

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 22, 2005, 18:59:02 ; Search time 188 Seconds
(without alignments)
23.371 Million cell updates/sec

Title: US-10-733-852-1

Perfect score: 55

Sequence: 1 WREPSPWALS 10

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: Genesecp21:*
2: Genesecp1980s:*
3: Genesecp1990s:*
4: Genesecp2000s:*
5: Genesecp2001s:*
6: Genesecp2002s:*
7: Genesecp2003as:*
8: Genesecp2003bs:*
9: Genesecp2004s:*
10: Genesecp2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	55	100.0	10	2	AAW81137
2	55	100.0	10	3	AAAY78912
3	55	100.0	10	3	AAAY53598
4	55	100.0	10	3	AAAY69642
5	55	100.0	10	4	AAAB68865
6	55	100.0	10	4	AAAB68988
7	55	100.0	10	8	ADU00345
8	55	100.0	10	9	ADV89959
9	55	100.0	16	2	AAW88511
10	55	100.0	16	2	AAW88510
11	55	100.0	16	3	AAAY53600
12	55	100.0	16	8	ADU00348
13	55	100.0	248	2	AAW89395
14	52	94.5	131	2	AAW08182
15	52	94.5	133	2	AAW08178
16	51	92.7	9	2	AAW08172
17	51	92.7	9	2	AAW84202
18	51	92.7	135	2	AAW08183
19	51	92.7	137	2	AAW08184
20	51	92.7	137	2	AAW08180
21	51	92.7	139	2	AAW08179
22	49	89.1	10	9	ADV89958
23	49	89.1	15	8	ADU00344
24	49	89.1	17	4	AAAB68989

25	45	81.8	10	9	ADV89960	Adv89960	Collagen
26	45	81.8	2813	2	AAW54347	Aaw54347	Canine vo
27	45	81.8	2813	3	AAAY70557	Aay70557	Canine vo
28	44	80.0	10	8	ADR59181	Adr59181	Collagen-
29	43	78.2	211	7	ABO60809	Abo60809	Klebsiell
30	40	72.7	219	3	AAG55414	Aag55414	Arabidops
31	40	72.7	990	3	ADY65948	Ady65948	S. mansoni
32	39	70.9	188	7	ABO69750	Abo69750	Pseudomon
33	38	69.1	112	6	ADA36063	Ada36063	Acinetoba
34	38	69.1	112	4	AAU57258	Aau57258	Propionib
35	38	69.1	112	6	ABM53777	Abm53777	Propionib
36	38	69.1	124	8	ADK47515	Adk47515	Streptoco
37	38	69.1	174	6	ABM64633	Abm64633	Propionib
38	38	69.1	391	8	AAAS43623	Aas43623	Bacterial
39	38	69.1	600	2	AAR86290	Aar86290	Protamino
40	38	69.1	600	8	ADJ87481	Adj87481	S. plymut
41	38	69.1	600	8	ADJ87457	Adj87457	S. plymut
42	38	69.1	600	8	ADJ87453	Adj87453	Protamino
43	38	69.1	629	2	AAR86287	Aar86287	Protamino
44	38	69.1	629	8	ADJ87447	Adj87447	Protamino
45	37	67.3	171	8	ADT49879	Adt49879	Human VWF

ALIGNMENTS

RESULT 1
AAW81137
ID AAW81137 standard; peptide; 10 AA.
XX
AC AAW81137;
XX
DT 05-MAR-1999 (first entry)
XX
DE Collagen-binding domain of von Willebrand Factor.
XX
KW Receptor binding region; ecotropic; gp70; retrovirus; envelope;
KW modified viral surface protein; moloney murine leukaemia virus;
KW extracellular matrix component; gene therapy; collagen.
XX
OS Homo sapiens.
XX
PN WO9844938-A1.
XX
PD 15-OCT-1998.
XX
PF 08-APR-1998; 98WO-US006936.
XX
PR 10-APR-1997; 97US-00837223.
XX
(UYSC-) UNIV SOUTHERN CALIFORNIA.
XX
Hall FL, Gordon EM, Anderson WF, Starnes VA;
WPI; 1998-594474/50.
N-PSDB; AAV68321.
XX
New vector particles targetted to extracellular matrix components -
contain modified viral protein including binding region for the
component, used for, e.g. gene therapy.
XX
Claim 6; Page 45; 73pp; English.
XX
The present invention describes a vector particle (A) which is targeted
to an extracellular matrix (ECM) component by modifying a viral surface
protein (VSP) to include a target polypeptide (TP) that includes a
binding region for the ECM component. The present sequence represents the
specifically claimed collagen-binding domain of von Willebrand Factor
from the present invention. (A) may also include a sequence encoding a
therapeutic agent (C), and are then used to express (C) in an animal
(gene therapy), specifically in cells located at a site (of injury) where
ECM components are exposed. Particular applications are in e.g.
cardiovascular disease, liver cirrhosis, thrombosis, vascular or

CC ulcerative lesions, inflammation, arthritis and tumours, also to promote
 CC wound healing. (A) may also be used in animal studies of gene therapy.
 CC Targeting to ECM concentrates (A) at sites of tissue injury, improving
 CC specificity and/or local concentration

SQ Sequence 10 AA;

Query Match 100.0%; Score 55; DB 2; Length 10;
 Best Local Similarity 100.0%; Pred. No. 0.0023;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WREPSFWALS 10
 |||||
 Db 1 WREPSFWALS 10

RESULT 2
 AAY78912
 ID AAY78912 standard; peptide; 10 AA.
 XX
 AC AAY78912;
 XX
 DT 23-MAY-2000 (first entry)
 XX
 DE Collagen binding domain (CBD) of von Willebrand factor (vWF).
 XX
 KW Collagen binding domain; von Willebrand factor; cardiovascular disease;
 KW angiogenesis modulatory agent; ulcerative lesion; inflammatory lesion;
 KW tumour; arthritis; myocardial infarction; peripheral artery disease;
 KW stroke; asthma; diabetic retinopathy; wound healing; tissue graft.

XX Bos sp.

XX W0200006195-A1.

XX 10-FEB-2000.

XX 30-JUL-1999; 99WO-US017297.

XX 31-JUL-1998; 98US-00127134.

XX (UYSC-) UNIV SOUTHERN CALIFORNIA.

XX Hall FL, Gordon EM, Starnes VA, Anderson WF;

XX WPI; 2000-195198/17.

XX New fusion polypeptides comprising a collagen binding domain and an
 PT angiogenesis modulating agent, useful for treating e.g. cardiovascular
 PT disease, ulcers, inflammation, tumors, arthritis or grafts.

XX Claim 4; Page 41; 63pp; English.

XX This sequence represents the collagen binding domain (CBD) peptide
 CC targets platelet aggregates to vascular lesions, and is involved in the
 CC recognition of exposed vascular collagen. The peptide is used in a fusion
 CC polypeptide which comprises a collagen binding domain and an angiogenesis
 CC modulating agent (e.g. vascular endothelial growth factor VEGF). The
 CC fusion polypeptides and nucleic acids can be used for locally altering
 CC circulation in a subject which may have a disorder such as cardiovascular
 CC disease, an ulcerative lesion, an inflammatory lesion, a tumour or
 CC arthritis. They can be used for treating myocardial infarction and
 CC peripheral artery disease, post-balloon angioplasty vascular stenosis,
 CC stroke, asthma, diabetic retinopathy or wound healing or genetic defects.
 CC They can also be used for preparing tissue grafts which can be used to
 CC repair or replace damaged tissue

SQ Sequence 10 AA;

Query Match 100.0%; Score 55; DB 3; Length 10;
 Best Local Similarity 100.0%; Pred. No. 0.0023;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WREPSFWALS 10
 |||||
 Db 1 WREPSFWALS 10

RESULT 3

AAY53598

ID AAY53598 standard; protein; 10 AA.

XX AAY53598;

XX 15-MAR-2000 (first entry)

XX Collagen binding domain from Von Willebrand factor.

XX Antitumour; anti-haemophilia; antidiabetic; anti-neurodegeneration;
 KW cardiovascular; retroviral vector; ecotropic receptor binding region;
 KW gp70; envelope; hypervariable; polypyrroline; gene therapy; tumour;
 KW severe combined immune deficiency; haemophilia; diabetes; animal model;
 KW Alzheimer's disease; Parkinson's disease; cystic fibrosis; recombinant;
 KW muscular dystrophy; cardiovascular disease; proteoliposome; ligand.

XX Homo sapiens.

XX W09955893-A1.

XX 04-NOV-1999.

XX 28-APR-1999; 99WO-IB000764.

XX 29-APR-1998; 98US-00069398.

XX (UYSC-) UNIV SOUTHERN CALIFORNIA.

XX Hall FL, Gordon EM, Anderson WF;

XX WPI; 2000-062038/05.

XX Retroviral vector containing both natural and modified envelope protein,
 PT to provide specific targeting and infectivity, useful in gene therapy.

XX Disclosure; Page 8; 41pp; English.

XX The invention relates to novel retroviral vector in which the ecotropic
 CC receptor binding region of the gp70 envelope protein (AAY53594) from a
 CC retroviral vector, preferably one based on the Moloney murine leukaemia
 CC virus (MuLV) is deleted or replaced by a non-retroviral sequence.
 CC Optionally the hypervariable polypyrroline region of the gp70 protein
 CC (AAY53595) may also be deleted or substituted. The modified retroviral
 CC vectors are used to introduce nucleic acid into eukaryotic cells,
 CC particularly for gene therapy of a very wide range of conditions, e.g.
 CC tumours, severe combined immune deficiency, haemophilia, diabetes,
 CC Alzheimer's and Parkinson's diseases, cystic fibrosis, muscular
 CC dystrophy, cardiovascular disease. The vectors can also be used to
 CC evaluate gene therapy methods in animal models, or to transduce cells for
 CC recombinant production of proteins. The envelope proteins are also useful
 CC as components of proteoliposomes for gene transfer or drug delivery. This
 CC peptide represents a collagen binding domain from von Willebrand factor.
 CC The peptide can be inserted into the retroviral vector in order to target
 CC the vector to cells expressing an extracellular matrix

SQ Sequence 10 AA;

Query Match 100.0%; Score 55; DB 3; Length 10;
 Best Local Similarity 100.0%; Pred. No. 0.0023;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WREPSFWALS 10
 |||||
 Db 1 WREPSFWALS 10

RESULT 4

AA69642
ID AAY69642 standard; peptide; 10 AA.

XX AC AAY69642;

XX DT 08-MAY-2000 (first entry)

XX DE Extracellular matrix targeting peptide #1.

XX KW Targeting peptide; membrane destabilisation; viral envelope protein;
KW gene therapy vector; extracellular matrix.

XX OS Unidentified.

XX PN WO200002909-A2.

XX PD 20-JAN-2000.

XX PF 08-JUL-1999; 99WO-IB001261.

XX PR 09-JUL-1998; 98US-00112544.

XX PA (UYSC-) UNIV SOUTHERN CALIFORNIA.

XX PI Rozenberg Y, Anderson WF;

XX PS WPI; 2000-147596/13.

XX CC New amphiphilic peptides used in developing modified enveloped viruses
CC for use in treating diseases associated with exposed extracellular matrix
CC component, in gene therapy.

XX PS Disclosure; Page 27; 83pp; English.

XX CC The invention relates to isolated peptides derived from the cytoplasmic
CC tail and/or the membrane-spanning region of a viral envelope protein.
CC Such peptides form an amphiphilic structure and have membrane
CC destabilising activity, facilitating the entry of viral particles into
CC cells and the efficient formation of viral particles. Accordingly, the
CC peptides may be attached to the membrane of an enveloped virus via its
CC incorporation (along with a targeting polypeptide) into an artificial
CC viral envelope protein. The membrane destabilising peptides of the
CC invention may be used for preparing a viral or synthetic vector for the
CC targeting and delivery of a therapeutic agent to an animal e.g., for
CC gene therapy, or for anti-tumour therapy. The amphiphilic peptides
CC reduce, inhibit or prevent the growth of a cell, virus, or virally
CC infected cell and therefore may be used as antibiotics, anti-viral
CC agents, or antimicrobial agents, and also as preservatives and
CC sterilants. Sequences AAY69639-Y69643 represent examples of targeting
CC peptides which may be used in the invention

XX SQ Sequence 10 AA;

Query Match 100.0%; Score 55; DB 3; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.0023;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WRPSFWALS 10

DB 1 WRPSFWALS 10

RESULT 5

AA68865
ID AAB68865 standard; peptide; 10 AA.

XX AC AAB68865;

XX DT 19-APR-2001 (first entry)

XX DE Von Willebrand factor collagen-binding domain decapeptide.

XX

KW Von Willebrand factor; collagen-binding domain; Osf2;
KW osteoblast specific factor 2; vulnery; muscular; gene therapy;
KW pre-mesenchymal; pre-haematopoietic stem cell; muscular dystrophy;
KW lipid storage disorder; skeletal disorder; bone marrow disorder;
KW gene therapy.

XX OS Unidentified.

XX PN WO200105944-A1.

XX PD 25-JAN-2001.

XX PF 20-JUL-2000; 2000WO-US019989.

XX PR 20-JUL-1999; 99US-0144786P.

XX PA (UYSC-) UNIV SOUTHERN CALIFORNIA.

XX PI Gordon EM, Hall FL;

XX DR WPI; 2001-147333/15.

XX CC Identifying pre-mesenchymal, pre-hematopoietic progenitor stem cells from
CC a population of cells, for use in treating e.g. a blood tissue related
CC disorder, involves identifying cells expressing osteoblast specific
CC factor 2.

XX PS Claim 11; Page 76; 92pp; English.

XX CC The present sequence is part of a fusion polypeptide which acts as a cell
CC proliferation-modulating agent in a method for identifying a pre-
CC mesenchymal, pre-haematopoietic stem cell from a population of cells. The
CC method comprises obtaining a population of cells from an animal species,
CC culturing the cells in vitro, contacting the cells with a cell
CC proliferation-modulating agent that induces Osf2 expression, and
CC identifying a pre-mesenchymal and pre-haematopoietic stem cell that
CC expresses Osf2. The identified cells are useful for ameliorating a
CC connective tissue-related disorder, blood tissue-related disorder or for
CC promoting bone marrow tissue regeneration. The blood tissue- and
CC connective tissue-related disorders are muscular dystrophy, lipid storage
CC disorders, skeletal disorders or bone marrow disorders. The cells are
CC also useful for promoting engraftment of mesenchymal and/or
CC haematopoietic stem cells in the recipient. The cells can also be used to
CC replace or supplement the corresponding cell type in a patient, or to
CC coat implants, thus acting as a barrier between the implant and the
CC patient. The cells may be genetically modified to prevent and/or treat
CC any conditions in which blood or connective tissue requires repair,
CC replacement or augmentation and are used in gene therapy techniques. The
CC present sequence is part of a fusion polypeptide which also comprises a
CC growth factor or an active fragment of a growth factor

XX SQ Sequence 10 AA;

Query Match 100.0%; Score 55; DB 4; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.0023;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WRPSFWALS 10

DB 1 WRPSFWALS 10

RESULT 6

AA68988
ID AAB68988 standard; peptide; 10 AA.

XX AC AAB68988;

XX DT 26-APR-2001 (first entry)

XX DE Von Willebrand factor-derived collagen-binding domain.

XX KW Collagen-binding domain; wound healing; epithelial tissue; EGF;

KW epidermal growth factor; cell proliferation; arthritis; osteoporosis;
 KW burn; tissue engineering; gene therapy.
 XX Unidentified.
 OS
 XX WO200107059-A1.
 PN
 XX 01-FEB-2001.
 PD
 XX 21-JUL-2000; 2000WO-US020055.
 PF
 XX 21-JUL-1999; 99US-0145488P.
 PR
 XX (UYSC-) UNIV SOUTHERN CALIFORNIA.
 PA
 XX Hall FL, Miami M, Gordon EM, Beart RW;
 PI
 XX WPI; 2001-182717/18.
 DR
 XX New epidermal growth factor collagen binding domain fusion polypeptides,
 PT useful for wound healing, promoting repair of colonic lesions, modulating
 PT epithelial cell proliferation or tissue regeneration.
 XX
 XX Claim 4; Page 63; 81pp; English.
 PS
 XX The present invention describes a fusion protein comprising a collagen-
 CC binding domain and an epithelial cell proliferation modulating agent such
 CC as epidermal growth factor (EGF). This can be used to target the agent to
 CC specific cells where it can then promote cell differentiation and
 CC encourage wound healing. This is useful not only in the promotion of
 CC wound healing following surgery, irradiation, infection and burning, but
 CC also in the treatment of arthritis, osteoporosis and other skeletal
 CC disorders, and in tissue engineering. Nucleic acids encoding the proteins
 CC can be used in gene therapy
 XX
 XX Sequence 10 AA;
 SQ
 Query Match 100.0%; Score 55; DB 4; Length 10;
 Best Local Similarity 100.0%; Pred. No. 0.0023;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 WREPSPFMAIS 10
 Db |||||
 1 WREPSPFMAIS 10
 RESULT 7
 ADU000345
 ID ADU000345 standard; peptide; 10 AA.
 AC
 AC ADU000345;
 XX
 XX 27-JAN-2005 (first entry)
 DT
 XX Truncated von Willebrand factor domain D2.
 XX
 XX D2 domain; von Willebrand factor; collagen; binding peptide;
 KW targeted delivery; vector; 4070A; amphotropic; envelope protein; viral;
 KW gag-pol; drug resistance; SV40; origin of replication;
 KW long terminal repeat; LTR; retroviral; packaging sequence; CMV; promoter;
 KW large T antigen; neoplastic disorder.
 XX
 XX Homo sapiens.
 OS Synthetic.
 OS
 XX WO2004093810-A2.
 PN
 XX 04-NOV-2004.
 PD
 XX 21-APR-2004; 2004WO-US012302.
 PF
 XX 21-APR-2003; 2003US-0464571P.
 PR
 XX
 XX

PA (BPEI-) EPEIUS BIOTECHNOLOGIES INC.
 XX Gordon EM, Hall FL;
 PI
 XX WPI; 2004-775863/76.
 DR
 XX Producing a targeted delivery vector for treating a neoplasm by culturing
 PT the producer cells transiently transfected with first, second and third
 PT plasmids, isolating and introducing the supernatant into a closed loop
 PT manifold system.
 XX
 XX Disclosure; Page 25; 103pp; English.
 PS
 XX This sequence represents a truncated D2 domain from von Willebrand
 CC factor. This peptide was used as a collagen binding peptide in the method
 CC of the invention. The method of the invention for producing a targeted
 CC delivery vector comprises: transiently transfecting a producer cell with
 CC first, second and third plasmids; culturing the producer cells under
 CC conditions that allow targeted delivery vector production and release
 CC into the supernatant of the culture; isolating and introducing the viral
 CC supernatant into a closed loop manifold system for collecting the viral
 CC particles; and collecting the targeted delivery vectors. The first
 CC plasmid comprises a nucleic acid sequence encoding the 4070A amphotrophic
 CC envelope protein modified to contain a collagen binding domain, where the
 CC nucleic acid sequence is operably linked to a promoter. The second
 CC plasmid comprises a nucleic acid sequence operably linked to a promoter
 CC and encoding a viral gag-pol polypeptide or a polypeptide that confers
 CC drug resistance on the producer cell, and an SV40 origin of replication.
 CC The third plasmid comprises a heterologous nucleic acid sequence operably
 CC linked to a promoter, where the sequence encodes a diagnostic or
 CC therapeutic polypeptide, 5' and 3' long terminal repeat sequences (LTR), a
 CC retroviral packaging sequence, a CMV promoter upstream of the 5' LTR, a
 CC nucleic acid sequence operably linked to a promoter, where the sequence
 CC encodes a polypeptide that confers drug resistance on the producer cell,
 CC an SV40 origin of replication, where the producer cell is a human cell,
 CC that expresses SV40 large T antigen. The method is useful in producing a
 CC targeted delivery vector for treating a neoplastic disorder.
 XX
 XX Sequence 10 AA;
 SQ
 Query Match 100.0%; Score 55; DB 8; Length 10;
 Best Local Similarity 100.0%; Pred. No. 0.0023;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 WREPSPFMAIS 10
 Db |||||
 1 WREPSPFMAIS 10
 RESULT 8
 ADV89959
 ID ADV89959 standard; peptide; 10 AA.
 XX
 AC ADV89959;
 XX
 XX 24-MAR-2005 (first entry)
 DT
 XX Collagen specific target binding moiety peptide seqid 10.
 DE
 XX magnetic resonance imaging; target binding moiety; TBM; collagen.
 KW
 XX Unidentified.
 OS
 XX WO2005001415-A2.
 PN
 XX 06-JAN-2005.
 PD
 XX 20-MAY-2004; 2004WO-US016029.
 PF
 XX 23-MAY-2003; 2003US-0473369P.
 PR
 XX (EPIX-) EPIX PHARM INC.
 PA
 XX

PI Amedio JC, Caravan PD, Jacques V, Zhou XL, Levy S;
 FI Kalageropoulos S, Greenfield M;
 XX WPI; 2005-101277/11.
 XX
 XX Composition useful for preparing targeted magnetic resonance imaging
 PT contrast agents for imaging targeted areas of the body e.g. thrombi and
 PT atherosclerotic lesions, comprises an organic chelating ligand or a metal
 PT chelate.
 XX
 XX Disclosure; SEQ ID NO 10; 82pp; English.
 PS
 XX The invention describes a composition (A) comprising an organic chelating
 CC ligand (1), where (A) has an enantiomeric excess of greater than 50% of
 CC an (R) or (S) isomer at the 2-position of (1). Also described are: a
 CC composition (B) comprising a metal chelate of formula (2) and its salt
 CC and having an enantiomeric excess of greater than 50% of an (R) or (S)
 CC isomer of the chelate at the 2 position; a magnetic resonance imaging
 CC (MRI) agent comprising (B) and optionally one or more Target Binding
 CC Moieties (TBW's); a composition (D) comprising an organic chelating
 CC ligand precursor of formula (3) and (4) where (D) has an enantiomeric
 CC excess of greater than 50% of an (R) or (S) isomer at the 2 position of
 CC the precursor; preparation of organic chelating ligand precursor; a
 CC method for converting an organic chelating ligand comprising one or more
 CC carboxylic acid esters to a metal chelate; a method for converting an
 CC organic chelating ligand precursor to a metal chelate; a composition (E)
 CC comprising an activated ester of an organic chelating ligand of formula
 CC (5) and (6), where (E) has an enantiomeric excess of greater than 50% of
 CC an (R) or (S) isomer at the 2 position of the chelating ligand; and a
 CC composition (F) comprising an activated ester of a metal chelate of
 CC formula (7) and its salt, where (F) has an enantiomeric excess of greater
 CC than 50% of an (R) or (S) isomer at the 2 position of the activated ester
 CC metal chelate. The organic chelating ligands and MRI contrast agents can
 CC be used to prepare targeted MRI contrast agents for imaging targeted
 CC areas of the body, including thrombi, atherosclerotic lesions,
 CC atherosclerotic plaque, the myocardium or the vasculature. The optical
 CC isomers of the organic chelating ligands can be synthesized in high yield
 CC and in high optical purity using inexpensive processes from readily
 CC available starting materials. This is the amino acid sequence of a
 CC collagen-binding TBM useful in the composition of the invention.
 XX
 XX Sequence 10 AA;
 SQ
 Query Match 100.0%; Score 55; DB 9; Length 10;
 Best Local Similarity 100.0%; Pred. No. 0.0023;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 WREPSFMALS 10
 DB |||||
 1 WREPSFMALS 10
 RESULT 9
 AAW88511
 ID AAW88511 standard; peptide; 16 AA.
 XX
 AC AAW88511;
 XX
 XX 12-MAR-1999 (first entry)
 DT
 XX Collagen binding domain insertion for pEA insert 2.
 DE
 XX Hypervariable proline region; amphotropic gp70 protein;
 KW human cytomegalovirus; hCMV; retrovirus; envelope protein.
 KW
 OS Synthetic.
 OS Human herpesvirus 5.
 XX
 XX WO9851700-A1.
 PN
 XX 19-NOV-1998.
 PD
 XX 13-MAY-1998; 98WO-US009718.
 PF

XX 14-MAY-1997; 97US-00856074.
 XX (UYSC-) UNIV SOUTHERN CALIFORNIA.
 PA
 XX Anderson WF, Wu BW;
 PI
 XX WPI; 1999-059696/05.
 DR
 XX New retroviral vector particle - has envelope polypeptide modified in the
 XX hypervariable proline region to enable cell targeting.
 PT
 XX Example 6; Page 38; 79pp; English.
 PS
 XX The present invention describes a retroviral vector particle (RVPI),
 CC having a modified retroviral envelope polypeptide which includes the
 CC hypervariable proline (HP) region modified to include a heterologous
 CC protein or peptide. Retroviral vectors made targetable by modification of
 CC the HP region enable the delivery of desired genes to a variety of cell
 CC types. Unlike prior art modifications in the receptor binding region, the
 CC particle does not disrupt envelope structure to impair folding,
 CC processing, and incorporation of the envelope protein. The present
 CC sequence represents a peptide from an example of the present invention
 XX
 XX Sequence 16 AA;
 SQ
 Query Match 100.0%; Score 55; DB 2; Length 16;
 Best Local Similarity 100.0%; Pred. No. 0.0037;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 WREPSFMALS 10
 DB |||||
 4 WREPSFMALS 13
 RESULT 10
 AAW88510
 ID AAW88510 standard; peptide; 16 AA.
 XX
 AC AAW88510;
 XX
 XX 12-MAR-1999 (first entry)
 DT
 XX Collagen binding domain insertion for pEA insert 1.
 DE
 XX Hypervariable proline region; amphotropic gp70 protein;
 KW human cytomegalovirus; hCMV; retrovirus; envelope protein.
 KW
 OS Synthetic.
 OS Human herpesvirus 5.
 XX
 XX WO9851700-A1.
 PN
 XX 19-NOV-1998.
 PD
 XX 13-MAY-1998; 98WO-US009718.
 PF
 XX 14-MAY-1997; 97US-00856074.
 XX
 XX (UYSC-) UNIV SOUTHERN CALIFORNIA.
 PA
 XX Anderson WF, Wu BW;
 PI
 XX WPI; 1999-059696/05.
 DR
 XX New retroviral vector particle - has envelope polypeptide modified in the
 XX hypervariable proline region to enable cell targeting.
 PT
 XX Example 6; Page 38; 79pp; English.
 PS
 XX The present invention describes a retroviral vector particle (RVPI),
 CC having a modified retroviral envelope polypeptide which includes the
 CC hypervariable proline (HP) region modified to include a heterologous

CC protein or peptide. Retroviral vectors made targetable by modification of
 CC the HP region enable the delivery of desired genes to a variety of cell
 CC types. Unlike prior art modifications in the receptor binding region, the
 CC particle does not disrupt envelope structure to impair folding,
 CC processing, and incorporation of the envelope protein. The present
 CC sequence represents a peptide from an example of the present invention
 XX
 XX Sequence 16 AA;
 SQ
 Query Match 100.0%; Score 55; DB 2; Length 16;
 Best Local Similarity 100.0%; Pred. No. 0.0037;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 WREPSFWALS 10
 Db 4 WREPSFWALS 13
 RESULT 11
 AAY53600
 ID AAY53600 standard; peptide; 16 AA.
 XX
 AC AAY53600;
 XX
 DT 15-MAR-2000 (first entry)
 XX
 DE Collagen binding domain peptide.
 XX
 KW Antitumour; anti-haemophilia; antidiabetic; anti-neurodegeneration;
 KW cardiovascular; retroviral vector; ecotropic receptor binding region;
 KW gp70; envelope; hypervariable; polypurine; gene therapy; tumour;
 KW severe combined immune deficiency; haemophilia; diabetes; animal model;
 KW Alzheimer's disease; Parkinson's disease; cystic fibrosis; recombinant;
 KW muscular dystrophy; cardiovascular disease; proteoliposome; ligand.
 XX
 OS Homo sapiens.
 XX
 PN WO955893-A1.
 XX
 PD 04-NOV-1999.
 XX
 PF 28-APR-1999; 99WO-18000764.
 XX
 PR 29-APR-1998; 98US-00069398.
 XX
 PA (UYSC-) UNIV SOUTHERN CALIFORNIA.
 XX
 PI Hall FL, Gordon EM, Anderson WF;
 XX
 PFPI; 2000-062038/05.
 XX
 DR
 XX
 PT Retroviral vector containing both natural and modified envelope protein,
 PT to provide specific targeting and infectivity, useful in gene therapy.
 XX
 PS Example 1; Page 16; 41pp; English.
 XX
 CC The invention relates to novel retroviral vector in which the ecotropic
 CC receptor binding region of the gp70 envelope protein (AAY53594) from a
 CC retroviral vector, preferably one based on the Moloney murine leukaemia
 CC virus (MMLV) is deleted or replaced by a non-retroviral sequence.
 CC Optionally the hypervariable polypurine region of the gp70 protein
 CC (AAY53595) may also be deleted or substituted. The modified retroviral
 CC vectors are used to introduce nucleic acid into eukaryotic cells, e.g.
 CC particularly for gene therapy of a very wide range of conditions, e.g.
 CC tumours, severe combined immune deficiency, haemophilia, diabetes,
 CC Alzheimer's and Parkinson's diseases, cystic fibrosis, muscular
 CC dystrophy, cardiovascular disease. The vectors can also be used to
 CC evaluate gene therapy methods in animal models, or to transduce cells for
 CC recombinant production of proteins. The envelope proteins are also useful
 CC as components of proteoliposomes for gene transfer or drug delivery. This
 CC peptide represents a collagen binding domain peptide which can be
 CC inserted into the retroviral vector in order to target the vector to
 CC cells expressing an extracellular matrix

XX
 SQ Sequence 16 AA;
 Query Match 100.0%; Score 55; DB 3; Length 16;
 Best Local Similarity 100.0%; Pred. No. 0.0037;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 WREPSFWALS 10
 Db 4 WREPSFWALS 13
 RESULT 12
 ADU00348
 ID ADU00348 standard; peptide; 16 AA.
 XX
 AC ADU00348;
 XX
 DT 27-JAN-2005 (first entry)
 XX
 DE Minimal collagen binding peptide.
 XX
 KW D2 domain; von Willebrand factor; collagen; binding peptide;
 KW targeted delivery; vector; 4070A; amphotrophic; envelope protein; viral;
 KW gag-pol; drug resistance; SV40; origin of replication;
 KW long terminal repeat; LTR; retroviral; packaging sequence; CMV; promoter;
 KW large T antigen; neoplastic disorder.
 XX
 OS Synthetic.
 XX
 PN WO2004093810-A2.
 XX
 PD 04-NOV-2004.
 XX
 PF 21-APR-2004; 2004WO-US012302.
 XX
 PR 21-APR-2003; 2003US-0464571P.
 XX
 PA (EPEI-) EPEUS BIOTECHNOLOGIES INC.
 XX
 PI Gordon EM, Hall FL;
 XX
 PFPI; 2004-775863/76.
 XX
 DR
 XX
 PT Producing a targeted delivery vector for treating a neoplasm by culturing
 PT the producer cells transiently transfected with first, second and third
 PT plasmids, isolating and introducing the supernatant into a closed loop
 PT manifold system.
 XX
 PS Example 1; Page 49; 103pp; English.
 XX
 CC This sequence represents a minimal collagen binding peptide. This peptide
 CC was used in the method of the invention for producing a targeted delivery
 CC vector. The method comprises: transiently transfecting a producer cell
 CC with first, second and third plasmids; culturing the producer cells under
 CC conditions that allow targeted delivery vector production and release
 CC into the supernatant of the culture; isolating and introducing the
 CC supernatant into a closed loop manifold system for collecting the viral
 CC particles; and collecting the targeted delivery vectors. The first
 CC plasmid comprises a nucleic acid sequence encoding the 4070A amphotrophic
 CC envelope protein modified to contain a collagen binding domain, where the
 CC nucleic acid sequence is operably linked to a promoter. The second
 CC plasmid comprises a nucleic acid sequence operably linked to a promoter
 CC and encoding a viral gag-pol polypeptide or a polypeptide that confers
 CC drug resistance on the producer cell, and an SV40 origin of replication.
 CC The third plasmid comprises a heterologous nucleic acid sequence operably
 CC linked to a promoter, where the sequence encodes a diagnostic or
 CC therapeutic polypeptide, 5' and 3' long terminal repeat sequences (LTR),
 CC a retroviral packaging sequence, a CMV promoter upstream of the 5' LTR, a
 CC nucleic acid sequence operably linked to a promoter, where the sequence
 CC encodes a polypeptide that confers drug resistance on the producer cell,
 CC an SV40 origin of replication, where the producer cell is a human cell
 CC that expresses SV40 large T antigen. The method is useful in producing a

CC purification of the fusion protein. The proteinase site is included to
 CC permit cleavage and release of the purification tag after purification if
 CC desired. The extracellular matrix binding site facilitates delivery of
 CC the fusion protein to the desired site of action. Delivery of the TGF-
 CC beta to the site to be treated reduces the amount of TGF-beta required to
 CC be administered to be effective and reduces the concentration of
 CC circulating TGF-beta which may result in undesirable effects. N.B. The
 CC present sequence is does not appear in the specification, but is derived
 CC from the sequences mentioned above
 XX
 SQ Sequence 131 AA;
 Query Match 94.5%; Score 52; DB 2; Length 131;
 Best Local Similarity 90.0%; Pred. No. 0.12;
 Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 WREPSFMALS 10
 DB 11 WREPSFMALA 20
 RESULT 15
 AAW08178
 ID AAW08178 standard; peptide; 133 AA.
 XX
 AC AAW08178;
 XX
 DT 26-AUG-1997 (first entry)
 DE TGF-beta fusion protein 22:15:20:30.
 DE Transforming growth factor-beta fusion protein; wound healing;
 KW artificial skin; surgery recovery time.
 XX
 OS Homo sapiens.
 OS Synthetic.
 XX
 FH Key Location/Qualifiers
 FT Region 1..6 /label= Purification_tag
 FT Region 7..12 /label= Proteinase_site
 FT Region 13..21 /label= Extracellular_matrix_binding_site
 FT Region 22..133 /label= TGF-beta1_active_fragment
 FT
 FT WO9639430-A1.
 XX
 PD 12-DEC-1996.
 XX
 PF 05-JUN-1996; 96WO-US0008973.
 XX
 PR 06-JUN-1995; 95US-00470837.
 XX
 PA (HALL/) HALL F L.
 PA (NIMNI/) NIMNI M E.
 PA (TUAN/) TUAN T.
 PA (WULL/) WU L.
 PA (CHEU/) CHEUNG D T.
 XX
 PI Hall FL, Nimni ME, Tuan T, Wu L, Cheung DT;
 XX
 DR WFI; 1997-043065/04.
 XX
 PT Prepn. of transforming growth factor-beta fusion protein - useful to
 PT reduce surgery recovery time and to prepare artificial skin.
 XX
 PS Example 1; Page; 59pp; English.
 XX
 CC A novel transforming growth factor-beta (TGF-beta) fusion protein
 CC comprises a purification tag and a TGF active fragment. Additionally, the
 CC fusion protein may comprise proteinase-sensitive linker sites and binding

CC domain so the protein sequence may contain some or all of the following
 CC elements: purification tag:proteinase site:ECM binding site: proteinase
 CC site:TGF-beta. The present sequence represents a fusion protein made up
 CC from a purification tag (AAW18225), a proteinase site (AAW08170), an
 CC extracellular matrix binding site (AAW08172) and a TGF active fragment
 CC (AAW08173). TGF-beta promotes wound healing, and the fusion protein can
 CC be used to reduce surgery recovery time and in the preparation of
 CC artificial skin. The inclusion of a purification tag facilitates
 CC purification of the fusion protein. The proteinase site is included to
 CC permit cleavage and release of the purification tag after purification if
 CC desired. The extracellular matrix binding site facilitates delivery of
 CC the fusion protein to the desired site of action. Delivery of the TGF-
 CC beta to the site to be treated reduces the amount of TGF-beta required to
 CC be administered to be effective and reduces the concentration of
 CC circulating TGF-beta which may result in undesirable effects. N.B. The
 CC present sequence is does not appear in the specification, but is derived
 CC from the sequences mentioned above
 XX

SQ Sequence 133 AA;

Query Match 94.5%; Score 52; DB 2; Length 133;
 Best Local Similarity 90.0%; Pred. No. 0.12;
 Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 WREPSFMALS 10
 DB 13 WREPSFMALA 22

Search completed: November 22, 2005, 19:14:51
 Job time : 191 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 22, 2005, 19:08:53 ; Search time 39 Seconds
(without alignments)
24.671 Million cell updates/sec

Title: US-10-733-852-1

Perfect score: 55

Sequence: 1 WREPSFWALS 10

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR 80:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	49	89.1	780	2 A34102	von Willebrand fac
2	39	70.9	326	2 C83263	hypothetical prote
3	38	69.1	48	2 T05927	heme A farnesyltra
4	38	69.1	318	2 AB2032	hypothetical prote
5	38	69.1	331	2 S48967	ethanolaminephosph
6	38	69.1	1150	2 T41260	hypothetical prote
7	37	67.3	317	2 S75694	hypothetical prote
8	37	67.3	2813	1 VWHU	von Willebrand fac
9	36	65.5	231	2 T04773	hypothetical prote
10	36	65.5	301	2 A97412	hypothetical prote
11	36	65.5	485	2 AF3106	glutamyl-tRNA synt
12	36	65.5	488	2 F98180	glutamyl-tRNA synt
13	36	65.5	693	1 S38427	NADPH-ferrihemopro
14	36	65.5	2469	2 H36812	hypothetical prote
15	36	65.5	2471	2 T42977	large tegument pro
16	36	65.5	2700	2 D88450	protein P21H11.2 [
17	36	65.5	5376	2 T42215	zonadhesin - mouse
18	35	63.6	278	2 T08252	probable transposa
19	35	63.6	434	2 E84879	probable heme A fa
20	35	63.6	469	2 T01579	heme A farnesyltra
21	35	63.6	736	2 A12162	hypothetical prote
22	35	63.6	1092	2 JX0312	differentiation-st
23	34	61.8	200	2 AC0788	cytochrome c-type
24	34	61.8	200	2 AC0369	cytochrome C-type
25	34	61.8	232	2 A83794	two-component resp
26	34	61.8	245	2 H87391	conserved hypothet
27	34	61.8	298	2 F84224	hypothetical prote
28	34	61.8	308	2 AB2960	hypothetical prote
29	34	61.8	312	2 B45335	probable heme A fa

30	34	61.8	312	2 H83976	cytochrome caa3 ox
31	34	61.8	319	2 A98323	ABC transporter, p
32	34	61.8	398	2 F75417	L-sorbose dehydr
33	34	61.8	409	2 H72210	argininosuccinate
34	34	61.8	409	2 C97376	probable transport
35	34	61.8	440	2 A12593	MPS permease [impo
36	34	61.8	485	2 G82037	potassium uptake p
37	34	61.8	683	2 T01442	hypothetical prote
38	34	61.8	861	2 S77086	hypothetical prote
39	34	61.8	1173	2 T30608	proteophosphoglyca
40	34	61.8	1530	2 S52239	brefeldin a resist
41	34	61.8	1530	2 T52010	hypothetical prote
42	34	61.8	1562	2 T43022	ATP-binding multidi
43	33	60.0	130	2 B72702	hypothetical prote
44	33	60.0	131	2 S24285	RNA-directed RNA p
45	33	60.0	168	2 PQ0481	1b protein - canin

ALIGNMENTS

RESULT 1

A34102

von Willebrand factor - bovine (fragments)

C/Species: Bos primigenius taurus (cattle)

C/Date: 30-Mar-1990 #sequence revision 05-May-1995 #text change 30-Jan-1998

C/Accession: A34102; S14768; A39540; S27196; S36600; S68346

R/Takagi, J.; Kasahara, K.; Sekiya, F.; Inada, Y.; Saito, Y.

J. Biol. Chem. 264, 10425-10430, 1989

A/Title: A collagen-binding glycoprotein from bovine platelets is identical to propolypepti

A/Reference number: A34102; MUID:89278101; PMID:2543668

A/Accession: A34102

A/Status: preliminary

A/Molecule type: protein

A/Residues: 1-21,78-98 <TAK>

A/Cross-references: UNIPARC:UPI0000177AD4; UNIPARC:UPI0000177ADS

R/Fujisawa, T.; Takagi, J.; Sekiya, F.; Goto, A.; Miake, F.; Saito, Y.

Eur. J. Biochem. 196, 673-677, 1991

A/Title: Monoclonal antibodies that inhibit binding of propolypeptide of von Willebrand

A/Reference number: S14768; MUID:91192039; PMID:1707363

A/Accession: S14768

A/Status: preliminary

A/Molecule type: protein

A/Residues: 22-51,52-77;112-149 <FUJ>

A/Cross-references: UNIPARC:UPI0000177AD6; UNIPARC:UPI0000177AD7; UNIPARC:UPI0000177AD8

R/Takagi, J.; Fujisawa, T.; Sekiya, F.; Saito, Y.

J. Biol. Chem. 266, 5575-5579, 1991

A/Title: Collagen-binding domain within bovine propolypeptide of von Willebrand factor.

A/Reference number: A39540; MUID:91170224; PMID:2005098

A/Accession: A39540

A/Status: preliminary

A/Molecule type: protein

A/Residues: 99-257,258-264 <TA2>

A/Cross-references: UNIPARC:UPI0000177AD9; UNIPARC:UPI0000177ADA

R/Bakshi, M.R.; Myers, J.C.; Howard, P.S.; Soprano, D.R.; Kirby, E.P.

Biochim. Biophys. Acta 1132, 325-328, 1992

A/Title: Sequencing of the primary adhesion domain of bovine von Willebrand factor.

A/Reference number: S27196; MUID:93041938; PMID:1420316

A/Accession: S27196

A/Molecule type: mRNA

A/Residues: 265-780 <BAK>

A/Cross-references: UNIPARC:UPI0000177ADB; EMBL:X63820

R/Bakshi, M.R.

submitted to the EMBL Data Library, January 1992

A/Reference number: S36600

A/Accession: S36600

A/Molecule type: mRNA

A/Residues: 265-615, 'SRVA', 620-687, 'M', 689-718, 'G', 720-780 <BA2>

A/Cross-references: UNIPARC:UPI0000177ADC; EMBL:X63820

R/Takagi, J.; Aoyama, T.; Ueki, S.; Ohba, H.; Saito, Y.; Lotand, L.

Eur. J. Biochem. 232, 773-777, 1995

A/Title: Identification of factor-XIIIa-reactive glutamyl residues in the propolypepti

A/Reference number: S68346; MUID:96028118; PMID:7588715

A;Accession: S68346
A;Molecule type: protein
A;Residues: 22-33; 'LSPVYAGRTCGLGNYN', 99-101, 'X', 103-122 <TAI>
A;Cross-references: UNIPARC:UPI0000177ADD; UNIPARC:UPI0000177ADE; UNIPARC:UPI0000177ADF
C;Superfamily: von Willebrand factor; von Willebrand factor type A repeat homology; von
C;Keywords: blood coagulation; cell binding; connective tissue; extracellular matrix; gl
F;513-681/Domain: von Willebrand factor type A repeat homology <VWA1>
F;755/Binding site: carboxydrate (Asn) (covalent) #status Predicted

Query Match 89.1%; Score 49; DB 2; Length 780;
Best Local Similarity 90.0%; Pred. No. 0.17;
Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 WREPSFWALS 10
DB 162 WREPSFWALS 171

RESULT 2
C83263
hypothetical protein PA3052 [imported] - Pseudomonas aeruginosa (strain PA01)
C;Species: Pseudomonas aeruginosa
C;Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004
C;Accession: C83263
R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; B
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,
; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A;Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho
A;Reference number: A82950; MUID:20437337; PMID:10984043
A;Accession: C83263
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-326 <STO>
A;Cross-references: UNIPROT:Q9HZF6; UNIPARC:UPI00000CS8EC; GB:AE004730; GB:AE004091; NID
A;Experimental source: strain PA01
C;Genetics:
A;Gene: PA3052

Query Match 70.9%; Score 39; DB 2; Length 326;
Best Local Similarity 66.7%; Pred. No. 6;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 WREPSFWAL 9
DB 265 WSEPTFMSL 273

RESULT 3
T05927
heme A farnesyltransferase homolog - barley (fragment)
C;Species: Hordeum vulgare (barley)
C;Date: 30-Apr-1999 #sequence_revision 30-Apr-1999 #text_change 09-Jul-2004
C;Accession: T05927
R;Hess, W.R.; Golz, R.R.; Boerner, T.
Plant Sci. 133, 191-201, 1998
A;Title: Analysis of randomly selected cDNAs reveals the expression of stress- and defe
A;Reference number: Z15411
A;Accession: T05927
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-48 <HES>
A;Cross-references: UNIPROT:O48635; UNIPARC:UPI00000A248F; EMBL:AJ222781; NID:e1203993;
A;Experimental source: cv. Haisa, leaf

Query Match 69.1%; Score 38; DB 2; Length 48;
Best Local Similarity 60.0%; Pred. No. 1.2;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 WREPSFWALS 10
DB 21 WQSPFWALA 30

RESULT 4

AB2032

hypothetical protein alr1808 [imported] - Nostoc sp. (strain PCC 7120)

C;Species: Nostoc sp. PCC 7120

A;Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120

C;Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Jul-2004

C;Accession: AB2032

R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi,

Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S.

DNA Res. 8, 205-213, 2001

A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana

A;Reference number: AB1807; MUID:21595285; PMID:11759840

A;Accession: AB2032

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-318 <KUR>

A;Cross-references: UNIPROT:Q8YW08; UNIPARC:UPI00000CE1FE; GB:BA000019; PIDN:BA073507.1;

A;Experimental source: strain PCC 7120

C;Genetics:

A;Gene: alr1808

Query Match 69.1%; Score 38; DB 2; Length 318;

Best Local Similarity 55.6%; Pred. No. 9.2;

Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 WREPSFWAL 9

DB 128 WQDPTFMSL 136

RESULT 5

S48967

ethanolaminephosphotransferase (EC 2.7.8.1) - yeast (Saccharomyces cerevisiae)

N;Alternate names: protein YHR123w; sn-1,2-diacylglycerol ethanolamine phosphotransferase

C;Species: Saccharomyces cerevisiae

C;Date: 02-Dec-1994 #sequence_revision 02-Dec-1994 #text_change 05-Oct-2004

C;Accession: S48967; S27296; A23714; S16700

R;Fulton, L.

submitted to the EMBL Data Library, June 1994

A;Description: The sequence of S. cerevisiae cosmid 9315.

A;Reference number: S48967

A;Accession: S48967

A;Molecule type: DNA

A;Residues: 1-391 <FUL>

A;Cross-references: UNIPROT:P22140; UNIPARC:UPI000012A0BD; EMBL:U10398; NID:G551328; PIDN:

R;Hjelmstad, R.H.; Bell, R.M.

submitted to the EMBL Data Library, April 1991

A;Reference number: S27296

A;Accession: S27296

A;Molecule type: DNA

A;Residues: 1-391 <HJE2>

A;Cross-references: UNIPARC:UPI000012A0BD; EMBL:M59311

R;Hjelmstad, R.H.; Bell, R.M.

J. Biol. Chem. 266, 5094-5103, 1991

A;Title: sn-1,2-diacylglycerol choline- and ethanolaminephosphotransferases in Saccharomy

A;Reference number: A23714; MUID:91161601; PMID:1848238

A;Accession: A23714

A;Molecule type: DNA

A;Residues: 1-95, 'V', 97, 'S', 99-277, 'FTSVDA TKL', 287-391 <HJE>

A;Cross-references: UNIPARC:UPI0000179595; GB:M59311

A;Note: the authors translated the codon GTA for residue 26 as Tyr

A;Note: the sequence from Fig. 7 is inconsistent with that from Fig. 2 in having 278-Leu,

C;Genetics:

A;Gene: SGD:EPT1

A;Cross-references: SGD:S0001165; MIPS:YHR123w

A;Map position: 8R

A;Introns: 17/2

C;Superfamily: choline/ethanolaminephosphotransferase

C;Keywords: transferase; transmembrane protein

F;49-65/Domain: transmembrane #status Predicted <TM1>

F;175-195/Domain: transmembrane #status Predicted <TM2>

F;216-236/Domain: transmembrane #status Predicted <TM3>

F/263-283/Domain: transmembrane #status predicted <TM4>
 F/289-309/Domain: transmembrane #status predicted <TM5>
 F/319-339/Domain: transmembrane #status predicted <TM6>
 F/347-367/Domain: transmembrane #status predicted <TM7>

Query Match 69.1%; Score 38; DB 2; Length 391;
 Best Local Similarity 60.0%; Pred. No. 11;
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 WRPSFMALS 10
 | : | | : | |
 Db 281 WMQPSFILTSL 290

RESULT 6
 T41260
 hypothetical protein SPCC285.14 - fission yeast (Schizosaccharomyces pombe)
 C/Species: Schizosaccharomyces pombe
 C/Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
 R/Seeger, K.; Harris, D.; Lyne, M.; Rajandream, M.A.; Barrell, B.G.
 submitted to the EMBL Data Library, September 1998
 A/Reference number: Z21981
 A/Accession: T41260
 A/Status: preliminary; translated from GB/EMBL/DBJ
 A/Molecule type: DNA
 A/Residues: 1-1150 <SEE>
 A/Cross-references: UNIPROT:O74501; UNIPARC:UPI0000134431; EMBL:AL031545; PIDN:CAA20853.
 A/Experimental source: strain 972h; cosmid c285
 C/Genetics:
 A/Gene: SPDB:SPCC285.14
 A/Map position: 3
 A/Introns: 10/1; 85/3; 642/3; 784/3

Query Match 69.1%; Score 38; DB 2; Length 1150;
 Best Local Similarity 100.0%; Pred. No. 37;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WRPSPF 6
 | | | | |
 Db 418 WRPSPF 423

RESULT 7
 S75694
 hypothetical protein slr1478 - Synechocystis sp. (strain PCC 6803)
 C/Species: Synechocystis sp.
 A/Variety: PCC 6803
 C/Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 09-Jul-2004
 C/Accession: S75694
 O. K.; Okumura, S.; Shimpō, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda
 R.; Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.;
 DNA Res. 3, 109-136, 1996
 A/Title: Sequence analysis of the genome of the unicellular cyanobacterium Synechocystis
 s.
 A/Reference number: S74322; MUID:97061201; PMID:8905231
 A/Accession: S75694
 A/Status: nucleic acid sequence not shown; translation not shown
 A/Molecule type: DNA
 A/Residues: 1-317 <KAN>
 A/Cross-references: UNIPROT:P74166; UNIPARC:UPI00000C1018; EMBL:D90912; GB:AB001339; NID
 A/Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996

Query Match 67.3%; Score 37; DB 2; Length 317;
 Best Local Similarity 55.6%; Pred. No. 14;
 Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 WRPSFMAL 9
 | : | | : | |
 Db 128 WRDPAFLEL 136

RESULT 8

VWUHU
 von Willebrand factor precursor - human
 C/Species: Homo sapiens (man)
 C/Date: 04-Dec-1986 #sequence_revision 30-Jun-1993 #text_change 09-Jul-2004
 C/Accession: A34480; S02377; A37139; S23676; A25366; A25369; A25366; S23618; S23645; A91044
 R/Mancuso, D.J.; Tuley, E.A.; Westfield, L.A.; Worrall, N.K.; Shelton-Inloes, B.B.; Sorace
 J. Biol. Chem. 264, 19514-19527, 1989
 A/Title: Structure of the gene for human von Willebrand factor.
 A/Reference number: A34480; MUID:90062044; PMID:2584182
 A/Accession: A34480
 A/Molecule type: DNA
 A/Residues: 1-2813 <MAN>
 A/Cross-references: UNIPROT:P04275; UNIPARC:UPI0000046823; EMBL:M25864
 R/Bonthron, D.; Orkin, S.H.
 Eur. J. Biochem. 171, 51-57, 1988
 A/Title: The human von Willebrand factor gene. Structure of the 5' region.
 A/Reference number: S02377; MUID:88111704; PMID:2828057
 A/Accession: S02377
 A/Molecule type: DNA
 A/Residues: 1-177 <BO2>
 A/Cross-references: UNIPARC:UPI000017431A; EMBL:X06828
 R/Mancuso, D.J.; Tuley, E.A.; Westfield, L.A.; Lester-Mancuso, T.L.; Le Beau, M.M.; Sorace
 Biochemistry 30, 253-269, 1991
 A/Reference number: S02377; MUID:88111704; PMID:2828057
 A/Title: Human von Willebrand factor gene and pseudogene: structural analysis and differ
 A/Reference number: A37139; MUID:91105089; PMID:1988024
 A/Accession: A37139
 A/Molecule type: DNA
 A/Residues: 990-1947 <WAD>
 A/Cross-references: UNIPARC:UPI000016B30B; GB:M60675; NID:G340357; PIDN:AAA61295.1; PID:
 A/Note: the authors translated the codon CGC for residue 156 as Gln
 R/Collins, C.J.; Underdahl, J.P.; Levene, R.B.; Ravera, C.P.; Morin, M.J.; Dombalagian, N.
 Proc. Natl. Acad. Sci. U.S.A. 84, 4393-4397, 1987
 A/Title: Molecular cloning of the human gene for von Willebrand factor and identification
 A/Reference number: S23676; MUID:87260814; PMID:3496594
 A/Accession: S23676
 A/Molecule type: DNA
 A/Residues: 2731-2813 <COL>
 A/Cross-references: UNIPARC:UPI000014240A; EMBL:M16945
 R/Bonthron, D.; Orr, E.C.; Mitscock, L.M.; Ginsburg, D.; Handin, R.I.; Orkin, S.H.
 Nucleic Acids Res. 14, 7125-7127, 1986
 A/Title: Nucleotide sequence of pre-pro-von Willebrand factor cDNA.
 A/Reference number: A25298; MUID:87016349; PMID:3489923
 A/Accession: A25298
 A/Molecule type: mRNA
 A/Residues: 1-470, 'V', 472-2813 <BON>
 A/Cross-references: UNIPARC:UPI000017431B; EMBL:X04385
 R/Verweij, C.L.; Diergaarde, P.J.; Hart, M.; Pannekoek, H.
 EMBO J. 5, 1839-1847, 1986
 A/Title: Full-length von Willebrand factor (vWF) cDNA encodes a highly repetitive protein
 A/Reference number: A91044; MUID:87004550; PMID:3019665
 A/Accession: A25469
 A/Molecule type: mRNA
 A/Residues: 1-470, 'V', 472-483, 'R', 485-1022, 'K', 1024-1025, 'E', 1027-1400 <VER>
 A/Cross-references: UNIPARC:UPI000017431C; EMBL:X04146
 A/Note: this sequence has been revised in reference A91056
 R/Verweij, C.L.; Diergaarde, P.J.; Hart, M.; Pannekoek, H.
 EMBO J. 5, 3074, 1986
 A/Reference number: A91056
 A/Accession: A25366
 A/Molecule type: mRNA
 A/Residues: 1021-1030 <VE2>
 A/Cross-references: UNIPARC:UPI000017431D
 A/Note: this is a revision to the sequence from reference A91044
 R/Shelton-Inloes, B.B.; Broze Jr. G.J.; Miletich, J.P.; Sadler, J.E.
 Biochem. Biophys. Res. Commun. 144, 657-665, 1987
 A/Title: Evolution of human von Willebrand factor: cDNA sequence polymorphisms, repeated
 A/Reference number: S23618; MUID:87213253; PMID:3495266
 A/Accession: S23618
 A/Molecule type: mRNA
 A/Residues: 1-120 <SH2>
 A/Cross-references: UNIPARC:UPI000016B30A; EMBL:M17588; NID:G799330; PIDN:AAA65940.1; PI
 A/Accession: S23645
 A/Molecule type: protein

RESULT 10
A97412
hypothetical protein agr_c_768 [imported] - Agrobacterium tumefaciens (strain C58, Cere
C/Species: Agrobacterium tumefaciens
C/Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 09-Jul-2004
C/Accession: A97412
R/Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldman,
A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.;
Science 294, 2323-2328, 2001
A/Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum
A/Reference number: A97359; MUID:21608551; PMID:11743194
A/Accession: A97412
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-301 <KUR>
A/Cross-references: UNIPROT:Q8U5L9; UNIPARC:UPI00000D2836; GB:AE007869; PIDN:AAK86250.1;
C/Genetics:
A/Gene: AGR_C_768
A/Map position: circular chromosome

Query Match 65.5%; Score 36; DB 2; Length 301;
Best Local Similarity 55.6%; Pred. No. 21;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
QY 1 WRPSFMAAL 9
||:|:|:
Db 99 WRDPTFTAV 107

RESULT 11
AF3106
glutamyl-tRNA synthetase gltX [imported] - Agrobacterium tumefaciens (strain C58, Dupont
C/Species: Agrobacterium tumefaciens
C/Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 18-Nov-2002
R/Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldman,
A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.;
Science 294, 2317-2323, 2001
A/Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,
star, E.W.
A/Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A/Reference number: AB2577; MUID:21608550; PMID:11743193
A/Accession: AF3106
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-485 <KUR>
A/Cross-references: UNIPARC:UPI0000164935; GB:AE008689; PIDN:AAI45268.1; PID:gl7742953;
A/Experimental source: strain C58 (Dupont)
C/Genetics:
A/Gene: gltX
A/Map position: linear chromosome
C/Superfamily: glutamate-tRNA ligase; glutamine-tRNA ligase homology

Query Match 65.5%; Score 36; DB 2; Length 485;
Best Local Similarity 60.0%; Pred. No. 36;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
QY 1 WRPSFMAAL 10
||:|:|:
Db 235 WQDPTFMHLS 244

RESULT 12
F98180
glutamyl-tRNA synthetase (glutamate-tRNA ligase) (glurs) [imported] - Agrobacterium tum
C/Species: Agrobacterium tumefaciens
C/Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 09-Jul-2004
C/Accession: F98180
R/Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldman,
A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.;

Science 294, 2323-2328, 2001
A/Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum
A/Reference number: A97359; MUID:21608551; PMID:11743194
A/Accession: F98180
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-488 <KUR>
A/Cross-references: UNIPROT:Q8U7H5; UNIPARC:UPI00000D25AA; GB:AE007870; PIDN:AAK88968.1;
C/Genetics:
A/Gene: AGR_L_791
A/Map position: linear chromosome
C/Superfamily: glutamate-tRNA ligase; glutamine-tRNA ligase homology
Query Match 65.5%; Score 36; DB 2; Length 488;
Best Local Similarity 60.0%; Pred. No. 36;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
QY 1 WRPSFMAAL 10
||:|:|:
Db 238 WQDPTFMHLS 247

RESULT 13
S38427
NADPH-ferrihemoprotein reductase (EC 1.6.2.4) - Aspergillus niger
N/Alternate names: NADPH-cytochrome P450 oxidoreductase
C/Species: Aspergillus niger
C/Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
C/Accession: S38427
R/van den Brink, J.; van Zeijl, C.; van den Hondel, C.; van Gorcom, R.
submitted to the EMBL Data Library, October 1993
A/Description: Cloning and characterization of the NADPH cytochrome P450 oxidoreductase
A/Reference number: S38427
A/Accession: S38427
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-693 <VAN>
A/Cross-references: UNIPROT:Q00141; UNIPARC:UPI000012FE4B; EMBL:Z26938; NID:G408110; PID:
C/Genetics:
A/Introns: 595/3
C/Superfamily: NADPH-ferrihemoprotein reductase; flavodoxin homology; NADPH-ferrihemopro
C/Keywords: flavoprotein; NADP; oxidoreductase
F/66-691/Domain: NADPH-ferrihemoprotein reductase homology <FEH>
F/68-220/Domain: flavodoxin homology <FLX>
Query Match 65.5%; Score 36; DB 1; Length 693;
Best Local Similarity 60.0%; Pred. No. 53;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
QY 1 WRPSFMAAL 10
||:|:|:
Db 215 WKPEPMWALS 224

RESULT 14
H36812
hypothetical protein ORF64 - saimiriine herpesvirus 1 (strain 11)
C/Species: saimiriine herpesvirus 1
A/Note: host Saimiri sciureus (common squirrel monkey)
C/Date: 16-Oct-1992 #sequence_revision 16-Oct-1992 #text_change 08-Oct-1999
C/Accession: H36812
R/Albrecht, J.
submitted to the EMBL Data Library, January 1992
A/Description: Primary structure of the herpesvirus saimiri genome.
A/Reference number: A36806
A/Accession: H36812
A/Molecule type: DNA
A/Residues: 1-2469 <ALB>
A/Cross-references: UNIPARC:UPI0000136BA0; GB:X64346; NID:G60320; PIDN:CAA45687.1; PID:G
R/Albrecht, J.C.; Nicholas, J.; Biller, D.; Cameron, K.R.; Biesinger, B.; Newman, C.; Wi
J. Virol. 66, 5047-5058, 1992
A/Title: Primary structure of the herpesvirus saimiri genome.
A/Reference number: A37309; MUID:92333688; PMID:1321287

A;Contents: annotation; protein-coding frames
A;Note: neither protein nor nucleotide sequence is given
C;Genetics:
A;Gene: 64

Query Match 65.5%; Score 36; DB 2; Length 2469;
Best Local Similarity 55.6%; Pred. No. 2.1e+02;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
QY 1 WREPSFMAL 9
| | | | |
Db 1796 WEHPSFLAI 1804

RESULT 15
T42977
large tegument protein - ateline herpesvirus 3 (strain 73)
C;Species: ateline herpesvirus 3
A;Variety: strain 73
C;Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 09-Jul-2004
C;Accession: T42977
R;Albrecht, J.C.; Fleckenstein, B.
submitted to the EMBL Data Library, August 1998
A;Description: Primary structure of the herpesvirus ateles genome.
A;Reference number: Z22274
A;Accession: T42977
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-2471 <ALB>
A;Cross-references: UNIPROT:Q9YTK3; UNIPARC:UPI00000EFC9B; EMBL:AF083424; PIDN:AAC95588.
A;Experimental source: strain 73

Query Match 65.5%; Score 36; DB 2; Length 2471;
Best Local Similarity 55.6%; Pred. No. 2.1e+02;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
QY 1 WREPSFMAL 9
| | | | |
Db 1796 WEHPSFMSI 1804

Search completed: November 22, 2005, 19:20:30
Job time : 42 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 22, 2005, 19:08:33 ; Search time 236 Seconds
(without alignments)
29.895 Million cell updates/sec

Title: US-10-733-852-1
Perfect score: 55
Sequence: 1 WREPSFMAIS 10

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Uniprot_05.80.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	49	89.1	937	1	VWF_BOVIN
2	45	81.8	2482	1	VWF_PIG
3	45	81.8	2813	1	VWF_CANFA
4	44	80.0	546	2	Q53MW3 ORYSA
5	41	74.5	2813	1	VWF_MOUSE
6	40	72.7	354	2	Q9FYA0 ARATH
7	40	72.7	1864	2	Q6VWH2_PACTO
8	39	70.9	326	2	Q9HZF6_PSEAE
9	39	70.9	577	2	Q6CED5_YARLI
10	39	70.9	1093	2	Q70535_RAT
11	38	69.1	48	2	O48635_HORVU
12	38	69.1	185	2	O5C3K8_SCHJA
13	38	69.1	318	2	Q8YW08_ANASP
14	38	69.1	371	2	Q7ZZP1_RANRI
15	38	69.1	371	2	Q800W0_RANDY
16	38	69.1	382	2	Q7NUM1_CHRYO
17	38	69.1	391	1	EPT1_YEAST
18	38	69.1	507	2	O415G1_GIBZE
19	38	69.1	552	2	Q51ZU2_MAGGR
20	38	69.1	556	2	Q7G761_ORYSA
21	38	69.1	556	2	Q941I2_ORYSA
22	38	69.1	968	2	Q7TUN3_PROWM
23	38	69.1	1150	1	YCRE_SCHPO
24	38	69.1	1427	2	Q4SS45_TETNG
25	37	67.3	87	2	Q4ZY32_PSEAS
26	37	67.3	252	2	Q7Q618_ANOGA
27	37	67.3	317	2	P74166_SYNY3
28	37	67.3	364	2	Q73X52_MYCPA
29	37	67.3	395	2	Q7V5M1_PROWM
30	37	67.3	403	2	Q603R7_METCA
31	37	67.3	403	2	Q4T081_TETNG

32 37 67.3 2282 1 ZAN_RABIT
33 37 67.3 2813 1 VWF_HUMAN
34 36 65.5 179 2 Q9LTV2_ARATH
35 36 65.5 199 2 Q5SR27_HUMAN
36 36 65.5 223 2 Q69QC1_ORYSA
37 36 65.5 231 2 Q9SYZ2_ARATH
38 36 65.5 276 2 Q4RJ07_TETNG
39 36 65.5 290 2 Q8VXY4_ARATH
40 36 65.5 301 2 Q8U5L9_AGR5
41 36 65.5 485 1 SYE_AGR5
42 36 65.5 627 2 Q6CWV3_KLULA
43 36 65.5 693 1 NCPR_A5BNG
44 36 65.5 695 2 Q5BFT5_EMENI
45 36 65.5 695 2 Q4WM67_ASPPU

P57999 oryctolagus
P04275 homo sapien
Q91TV2 arabidopsis
Q5SR27 homo sapien
Q69QC1 oryza sativ
Q9SYZ2 arabidopsis
Q4RJ07 tetraodon n
Q8VXY4 arabidopsis
Q8U5L9 agrobacteri
Q8U7H5 agrobacteri
Q6CWV3 kluyveromyc
Q00141 aspergillus
Q5BFT5 aspergillus
Q4WM67 aspergillus

ALIGNMENTS

RESULT 1
VWF_BOVIN
ID VWF_BOVIN STANDARD; PRT; 937 AA.
AC P80012; Q28011;
DT 01-NOV-1995 (Rel. 32, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE Von Willebrand factor precursor (VWF) (Fragment).
GN Name=VWF; Synonym=F8VWF;
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;
OC Pecora; Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=97307751; PubMed=9165093; DOI=10.1016/S0167-4838(97)00043-5;
RA Janel N., Ribba A.S., Cherel G., Kerbirou-Nabias D., Meyer D.;
RT "Primary structure of the propeptide and factor VIII-binding domain of
RT bovine von Willebrand factor.";
RL Biochim. Biophys. Acta 1339:4-8(1997).
RN [2]
RP NUCLEOTIDE SEQUENCE OF 1-177.
RX MEDLINE=96144290; PubMed=8566794; DOI=10.1016/0378-1119(95)00696-6;
RA Janel N., Schwachtgen J.L., Bakhshi M.R., Barek L., Meyer D.,
RA Kerbirou-Nabias D.;
RT "Comparison of the 5'-flanking sequences of the human and bovine von
RT Willebrand factor-encoding genes reveals alternation of highly
RT homologous domains with species-specific Alu-type repeats.";
RL Gene 167:291-295(1995).
RN [3]
RP PARTIAL PROTEIN SEQUENCE.
RX MEDLINE=91192039; PubMed=1707363;
RA Fujisawa T., Takagi J., Sekiya F., Goto A., Miake F., Saito Y.;
RT "Monoclonal antibodies that inhibit binding of propolypeptide of von
RT Willebrand factor to collagen. Localization of epitopes.";
RL Eur. J. Biochem. 196:673-677(1991).
RN [4]
RP PARTIAL PROTEIN SEQUENCE.
RX MEDLINE=96028118; PubMed=7588715;
RA Takagi J., Aoyama T., Ueki S., Ohba H., Saito Y., Lorand L.;
RT "Identification of factor-XIIIa-reactive glutamyl residues in the
RT propolypeptide of bovine von Willebrand factor.";
RL Eur. J. Biochem. 232:773-777(1995).
RN [5]
RP REVIEW.
RX PubMed=12871266; DOI=10.1046/j.1538-7836.2003.00260.x;
RA Ruggeri Z.M.;
RT "Von Willebrand factor, platelets and endothelial cell interactions.";
RL J. Thromb. Haemost. 1:1335-1342(2003).
CC -!- FUNCTION: Important in the maintenance of hemostasis, it promotes
CC adhesion of platelets to the sites of vascular injury by forming a
CC molecule bridge between sub-endothelial collagen matrix and
CC platelet-surface receptor complex GPIb-IX-V. Also acts as a

CC chaperone for coagulation factor VIII, delivering it to the site
CC of injury, stabilizing its heterodimeric structure and protecting
CC it from premature clearance from plasma.
CC -1- SUBUNIT: Multimeric (By similarity).
CC -1- SUBCELLULAR LOCATION: Secreted. Localized to storage granules (By
CC similarity).
CC -1- DOMAIN: The propeptide is required for multimerization of VWF and
CC for its targeting to storage granules (By similarity).
CC -1- PTM: All cysteine residues are involved in intrachain or
CC interchain disulfide bonds (By similarity).
CC -1- SIMILARITY: Contains at least 3 VWFD domains.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
DR EMBL; Y09353; CAA70525.1; -; mRNA.
DR EMBL; U28147; AAA96953.1; -; mRNA.
DR InterPro; IPR006207; Cys_knot_C.
DR InterPro; IPR002919; Prot_Inh_CR_TIL.
DR InterPro; IPR001007; VWF_C.
DR InterPro; IPR001846; VWF_D.
DR Pfam; PF01826; TIL; 2.
DR Pfam; PF00094; VWD; 2.
DR SMART; SM00214; VWD; 1.
DR SMART; SM00216; VWD; 2.
DR PROSITE; PS01185; CTCK_1; PARTIAL.
DR PROSITE; PS01225; CTCK_2; PARTIAL.
DR PROSITE; PS01208; WFC_1; PARTIAL.
KW Blood coagulation; Cell adhesion; Direct protein sequencing;
KW Extracellular matrix; Glycoprotein; Hemostasis; Repeat; Signal.
FT SIGNAL 1 22 Potential.
FT PROPEP 23 762 By similarity.
FT CHAIN 763 >937 Von Willebrand factor.
FT DOMAIN 35 179 VWFD 1.
FT DOMAIN 387 540 VWFD 2.
FT DOMAIN 841 >937 VWFD 3.
FT REGION 787 832 El.
FT REGION 825 852 Amino-terminal.
FT SITE 410 410 CX.
FT SITE 414 414 Factor XIIIa-binding.
FT SITE 605 605 Factor XIIIa-binding.
FT SITE 605 605 Factor XIIIa-binding.
FT CARBOHYD 156 156 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 211 211 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 665 665 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 856 856 N-linked (GlcNAc...) (Potential).
FT DISULFID 766 807 By similarity.
FT DISULFID 775 803 By similarity.
FT DISULFID 913 920 By similarity.
FT CONFLICT 330 330 Missing (in Ref. 3).
FT CONFLICT 523 523 Missing (in Ref. 4).
FT CONFLICT 528 528 G -> Q (in Ref. 4).
FT NON_TER 937 937
SQ SEQUENCE 937 AA; 102599 MW; 9BF4C94A254A5629 CRC64;
Query Match 89.1%; Score 49; DB 1; Length 937;
Best Local Similarity 90.0%; Pred. No. 1.5;
Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 WREPSFWALS 10
DB 641 WREPSFWALS 650
|||||
RESULT 2
VWF_PIG STANDARD; PRT; 2482 AA.
AC Q28833;
DT 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE Von Willebrand factor precursor (VWF) (Fragment).
GN Names=VWF; Synonyms=F8VWF;
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Suina; Suidae;
OC Sus
OX NCBI_TaxID=9823;
OX [1]
RN NUCLEOTIDE SEQUENCE.
RP Seaman W.T., Read M.S., Bellinger D.A., Nichols T.C.;
RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
RN NUCLEOTIDE SEQUENCE OF 397-553.
RX MEDLINE=93356762; PubMed=8352759;
RA Lavergne J.M., Piao Y.C., Ferreira V., Kerbirou-Nabias D.,
RA Bahnak B.R., Meyer D.;
RT "Primary structure of the factor VIII binding domain of human, porcine
RT and rabbit von Willebrand factor";
RL Biochem. Biophys. Res. Commun. 194:1019-1024 (1993).
RN [3]
RP REVIEW.
RX PubMed=12871266; DOI=10.1046/j.1538-7836.2003.00260.x;
RA Ruggeri Z.M.;
RT "Von Willebrand factor, platelets and endothelial cell interactions.";
RL J. Thromb. Haemost. 1:1335-1342 (2003).
CC -1- FUNCTION: Important in the maintenance of hemostasis, it promotes
CC adhesion of platelets to the sites of vascular injury by forming a
CC molecular bridge between sub-endothelial collagen matrix and
CC platelet-surface receptor complex, glycoprotein Ib/alpha/IX/V. Also
CC acts as a chaperone for coagulation factor VIII, delivering it to
CC the site of injury, stabilizing its heterodimeric structure and
CC protecting it from premature clearance from plasma (By
CC similarity).
CC -1- SUBUNIT: Multimeric (By similarity).
CC -1- SUBCELLULAR LOCATION: Secreted. Localized to storage granules (By
CC similarity).
CC -1- DOMAIN: The propeptide is required for multimerization of VWF and
CC for its targeting to storage granules (By similarity).
CC -1- PTM: All cysteine residues are involved in intrachain or
CC interchain disulfide bonds (By similarity).
CC -1- SIMILARITY: Contains 1 CTCK (C-terminal cysteine knot-like) domain.
CC -1- SIMILARITY: Contains 3 WFA domains.
CC -1- SIMILARITY: Contains 3 WFC domains.
CC -1- SIMILARITY: Contains at least 1 VWFD domain.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
DR EMBL; AF052036; AAC06229.1; -; mRNA.
DR EMBL; S64541; AAB27829.2; -; mRNA.
DR PIR; P05563; P05563.
DR HSSP; P04275; 1ATZ.
DR SNR; Q28833; 931-1138, 1353-1542.
DR InterPro; IPR006207; Cys_knot_C.
DR InterPro; IPR006210; IEFG.
DR InterPro; IPR002919; Prot_Inh_CR_TIL.
DR InterPro; IPR008037; Prot_inh_PMP.
DR InterPro; IPR002035; VWF_A.
DR InterPro; IPR001007; VWF_C.
DR InterPro; IPR001846; VWF_D.
DR Pfam; PF05375; Pacifastin_1; 1.
DR Pfam; PF01826; TIL; 3.
DR Pfam; PF00092; VWA; 3.
DR Pfam; PF00093; WVC; 3.
DR Pfam; PF00094; VWD; 3.
DR PRINTS; PR00453; VWFADOMAIN.
DR SMART; SM00041; CT; 1.
DR SMART; SM00181; EGF; 1.

DR SMART; SM00327; WVA; 3.
 DR SMART; SM00214; WVC; 5.
 DR SMART; SM00216; WVD; 3.
 DR PROSITE; PS01185; CTCK_1; 1.
 DR PROSITE; PS01225; CTCK_2; 1.
 DR PROSITE; PS50234; WVA; 3.
 DR PROSITE; PS01208; WVC; 1; 3.
 DR PROSITE; PS0184; WVC; 2; 3.
 KW Blood coagulation; Cell adhesion; Extracellular matrix; Glycoprotein;
 HM Hemostasis; Repeat.
 FT PROPEP <1 437 By similarity.
 FT CHAIN 438 2482 Von Willebrand factor.
 FT DOMAIN 62 215 WFWD 2.
 FT DOMAIN 541 687 WFWD 3.
 FT DOMAIN 947 1127 WFVA 1.
 FT DOMAIN 1167 1334 WFVA 2.
 FT DOMAIN 1360 1540 WFVA 3.
 FT DOMAIN 1619 1771 WFWD 4.
 FT DOMAIN 1924 1997 WFWD 1.
 FT DOMAIN 2098 2164 WFWD 2.
 FT DOMAIN 2249 2314 WFWD 3.
 FT DOMAIN 2393 2481 CTCK.
 FT REGION 438 461 Amino-terminal.
 FT REGION 462 507 E1.
 FT REGION 500 527 E2.
 FT MOTIF 2176 2178 Cell attachment site (Potential).
 FT CARBOHYD 340 340 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 905 905 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 1184 1184 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 1243 1243 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 1640 1640 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 1892 1892 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 1959 1959 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 2026 2026 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 2069 2069 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 2215 2215 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 2254 2254 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 2459 2459 N-linked (GlcNAc...) (Potential).
 FT DISULFID 441 482 By similarity.
 FT DISULFID 450 478 By similarity.
 FT DISULFID 484 495 By similarity.
 FT DISULFID 541 670 By similarity.
 FT DISULFID 563 705 By similarity.
 FT DISULFID 572 667 By similarity.
 FT DISULFID 588 595 By similarity.
 FT DISULFID 734 758 By similarity.
 FT DISULFID 745 785 By similarity.
 FT DISULFID 763 765 By similarity.
 FT DISULFID 800 804 By similarity.
 FT DISULFID 823 843 By similarity.
 FT DISULFID 827 839 By similarity.
 FT DISULFID 870 873 By similarity.
 FT DISULFID 908 911 By similarity.
 FT DISULFID 942 1128 By similarity.
 FT DISULFID 1338 1339 By similarity.
 FT DISULFID 1355 1541 By similarity.
 FT DISULFID 1548 1573 By similarity.
 FT DISULFID 1568 1609 Or C-1568 with C-1611 (By similarity).
 FT DISULFID 1596 1757 By similarity.
 FT DISULFID 1619 1754 By similarity.
 FT DISULFID 1641 1792 By similarity.
 FT DISULFID 1662 1670 By similarity.
 FT DISULFID 2393 2443 By similarity.
 FT DISULFID 2408 2457 By similarity.
 FT DISULFID 2419 2473 By similarity.
 FT DISULFID 2423 2475 By similarity.
 FT DISULFID ? 2480 By similarity.
 FT NON TER 1 1
 SQ SEQUENCE "2482 AA; 272397 MW; D499B7DFBBCAEDD CRC64;
 Query Match 81.8%; Score 45; DB 1; Length 2482;
 Best Local Similarity 80.0%; Pred. No. 25;

Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 WREPFWALS 10
 DB 316 WREPFGFALS 325
 RESULT 3
 ID VWF CANFA STANDARD; PRT; 2813 AA.
 AC Q28295; Q28311; Q2TS14;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 10-MAY-2005 (Rel. 47, Last annotation update)
 DE Von Willebrand factor precursor (VWF).
 GN Name=VWF; Synonyms=F8VWF;
 OS Canis familiaris (Dog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Laurasiatheria; Carnivora; Fissipedia; Canidae;
 OC Canis.
 OX NCBI_TaxID=9615;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RA Stoy S.J., Shibuya H., Nonneman D.J., Holzhauer J., Mohammed I.H.,
 RA Johnson G.S.;
 RL Submitted (AUG-1996) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP NUCLEOTIDE SEQUENCE, SUBUNIT, AND SUBCELLULAR LOCATION.
 RC TISSUE=Heart;
 RX PubMed=10961880;
 RA Haberichter S.L., Fahs S.A., Montgomery R.R.;
 RT "von Willebrand factor storage and multimerization: 2 independent
 RT intracellular processes.";
 RL Blood 96:1808-1815(2000).
 RN [3]
 RP SEQUENCE REVISION TO 55.
 RA Montgomery R.R., Fahs S., Montgomery M.W.;
 RL Submitted (JAN-2004) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP NUCLEOTIDE SEQUENCE, AND DISEASE.
 RC STRAIN=Scottish terrier; TISSUE=Uterus;
 RX PubMed=10668811; DOI=10.1892/0891-6640(2000)014<0010:MCVWDI>2.3.CO;2;
 RA Venta P.J., Li J., Yuzbasiyan-Gurkan V., Brewer G.J., Schall W.D.;
 RT "Mutation causing von Willebrand's disease in Scottish Terriers.";
 RL J. Vet. Intern. Med. 14:10-19(2000).
 RN [5]
 RP NUCLEOTIDE SEQUENCE OF 1234-1669.
 RC TISSUE=Blood;
 RA Mancuso D.J., Christopherson P.A., Kroner P.A., Montgomery R.R.;
 RT "The canine von Willebrand factor gene: sequence and expression of a
 RT region encoding the glycoprotein Ib/IX binding domain.";
 RL Submitted (JAN-1994) to the EMBL/GenBank/DBJ databases.
 RN [6]
 RP REVIEW.
 RX PubMed=12871266; DOI=10.1046/j.1538-7836.2003.00260.x;
 RA Ruggeri Z.M.;
 RT "Von Willebrand factor, platelets and endothelial cell interactions.";
 RL J. Thromb. Haemost. 1:1335-1342(2003).
 CC -I- FUNCTION: Important in the maintenance of hemostasis, it promotes
 CC adhesion of platelets to the sites of vascular injury by forming a
 CC molecular bridge between sub-endothelial collagen matrix and
 CC platelet-surface receptor complex, glycoprotein IbaIpha/IX/V. Also
 CC acts as a chaperone for coagulation factor VIII, delivering it to
 CC the site of injury, stabilizing its heterodimeric structure and
 CC protecting it from premature clearance from plasma (By
 CC similarity).
 CC -I- SUBUNIT: Multimeric.
 CC -I- SUBCELLULAR LOCATION: Secreted. Localized to storage granules.
 CC -I- TISSUE SPECIFICITY: Blood.
 CC -I- DOMAIN: The propeptide is required for multimerization of VWF and
 CC for its targeting to storage granules.
 CC -I- PTM: All cysteine residues are involved in intrachain or
 CC interchain disulfide bonds (By similarity).

RA Buell C., Yuan Q., Ouyang S., Liu J., Wang A., Maiti R., Lin H.,
 RA Zhu W., Hamilton J., Jones K., Tallon L., Feldblyum T., Tsitrin T.,
 RA Bera J., Kim M., Jin S., Fadrosch D., Vuong H., Overton II L.,
 RA Reardon M., Weaver B., Johri S., Lewis M., Utterback T., Van Aken S.,
 RA Wortman J., Haas B., Koo H., Ziemann V., Hsiao J., Iobst S.,
 RA de Vazeilles A., White O., Salzberg S., Fraser C.,
 RL Submitted (OCT-2002) to the ENBL/GenBank/DBJ databases.
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RA Buell R.;
 RL Submitted (APR-2005) to the ENBL/GenBank/DBJ databases.
 DR ENBL: AC135561; AAX96628.L.; -: Genomic DNA.
 SQ SEQUENCE 546 AA; 59642 MW; 8F96C80977A85839 CRC64;
 Query Match 80.0%; Score 44; DB 2; Length 546;
 Best Local Similarity 70.0%; Pred. No. 8.2;
 Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
 QY 1 WREPSFNALS 10
 Db 458 WREPVFMAIT 467
 RESULT 5
 VWF_MOUSE
 ID_VWF_MOUSE STANDARD; PRT; 2813 AA.
 AC Q8CIZ8; Q60863; Q6XUV6; Q8BIU9; Q8CGN0; Q9JK16;
 DT 01-FEB-2005 (Rel. 46, Created)
 DT 01-FEB-2005 (Rel. 46, Last sequence update)
 DT 13-SEP-2005 (Rel. 48, Last annotation update)
 DE Von Willebrand factor precursor (VWF) [Contains: Von Willebrand
 DE antigen II].
 GN Name=Vwf;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
 OC Muridea; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP NUCLEOTIDE SEQUENCE (ISOFORM 1).
 RC STRAIN=C57BL/6; TISSUE=Lung;
 RA Chitta M.S., Duhe R.J., Kermod J.C.;
 RT "Cloning of full-length murine von Willebrand factor cDNA";
 RL Submitted (DEC-2002) to the ENBL/GenBank/DBJ databases.
 RN [2]
 RP NUCLEOTIDE SEQUENCE (ISOFORM 1).
 RC STRAIN=BALB/c;
 RA Lenting P.J., Westein E., de Groot P.G., Denis C.V.;
 RT "Murine von Willebrand factor";
 RL Submitted (JUL-2003) to the ENBL/GenBank/DBJ databases.
 RN [3]
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] (ISOFORM 2).
 RC STRAIN=C57BL/6J; TISSUE=Hippocampus;
 RX MEDLINE=22354683; PubMed=12466851; DOI=10.1038/nature01266;
 RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,
 RA Nkaido I., Osato N., Saio R., Suzuki H., Yamanaka I., Kiyosawa H.,
 RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,
 RA Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
 RA Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,
 RA Blake J.A., Bradt D., Brusci V., Chothia C., Corbani L.E., Cousins S.,
 RA Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,
 RA Gaasterland T., Gariboldi M., Glessi C., Godzik A., Gough J.,
 RA Grimmond S., Guetinch S., Hirokawa N., Jackson I.J., Jarvis E.D.,
 RA Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L.,
 RA Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
 RA Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,
 RA Negashima T., Numata K., Okido T., Pavan W.J., Perlea G., Pesole G.,
 RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,
 RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
 RA Sadelin A., Schneider C., Sempke C.A., Setou M., Shimada K.,
 RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
 RA Varadar R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,
 RA Wilming L.G., Wynshaw-Boris A., Yanagisawa M., Yang I., Yang L.,

RA Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
 RA Hiroane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,
 RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,
 RA Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,
 RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,
 RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
 RA Birney E., Hayashizaki Y.;
 RT "Analysis of the mouse transcriptome based on functional annotation of
 RT 60,770 full-length cDNAs";
 RL Nature 420:563-573(2002).
 RN [4]
 RP NUCLEOTIDE SEQUENCE OF 1238-1658 (ISOFORM 1).
 RX MEDLINE=20185297; PubMed=10722222; DOI=10.1098/rspb.2000.1014;
 RA Ruchon D., Catzeffis F.M., Douzery E.J.P.;
 RT "Variation of molecular datings, evolution of rodents and the
 RT phylogenetic affinities between Ctenodactylidae and Hystriognathi";
 RL Proc. R. Soc. Lond., B, Biol. Sci. 267:393-402(2000).
 RN [5]
 RP NUCLEOTIDE SEQUENCE OF 1298-1684 (ISOFORM 1).
 RC STRAIN=BALB/c;
 RA MEDLINE=94250904; PubMed=8193357;
 RA Nichols W.C., Cooney K.A., Mohlke K.L., Ballew J.D., Yang A.,
 RA Bruck M.E., Redington M., Novak E.K., Swank R.T., Ginsburg D.;
 RT "von Willebrand disease in the RIIIS/J mouse is caused by a defect
 RT outside of the von Willebrand factor gene";
 RL Blood 83:3225-3231(1994).
 RN [6]
 RP ERRATUM.
 RX PubMed=7662996;
 RA Nichols W.C., Cooney K.A., Mohlke K.L., Ballew J.D., Yang A.,
 RA Bruck M.E., Redington M., Novak E.K., Swank R.T., Ginsburg D.;
 RL Blood 86:2461-2461(1995).
 RN [7]
 RP MUTAGENESIS OF ARG-1205, AND TISSUE SPECIFICITY.
 RC STRAIN=C57BL/6J;
 RX PubMed=14613933; DOI=10.1074/jbc.M10436200;
 RA Lenting P.J., Westein E., Terraube V., Ribba A.-S., Huizinga E.G.,
 RA Meyer D., de Groot P.G., Denis C.V.;
 RT "An experimental model to study the in vivo survival of von Willebrand
 RT factor. Basic aspects and application to the R1205H mutation";
 RL J. Biol. Chem. 279:12102-12109(2004).
 RN [8]
 RP REVIEW.
 RX PubMed=12871266; DOI=10.1046/j.1538-7836.2003.00260.x;
 RA Ruggeri Z.M.;
 RT "Von Willebrand factor, platelets and endothelial cell interactions";
 RL J. Thromb. Haemost. 1:1335-1342(2003).
 CC -1- FUNCTION: Important in the maintenance of hemostasis, it promotes
 CC adhesion of platelets to the sites of vascular injury by forming a
 CC molecular bridge between sub-endothelial collagen matrix and
 CC platelet-surface receptor complex GPIIb-IX-V. Also acts as a
 CC chaperone for coagulation factor VIII, delivering it to the site
 CC of injury, stabilizing its heterodimeric structure and protecting
 CC it from premature clearance from plasma (By similarity).
 CC -1- SUBUNIT: Multimeric (by similarity).
 CC -1- SUBCELLULAR LOCATION: Secreted. Localized to storage granules (By
 CC similarity).
 CC -1- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=2;
 CC Name=1;
 CC IsoId=Q8CIZ8-1; Sequence=Displayed;
 CC Name=2;
 CC IsoId=Q8CIZ8-2; Sequence=VSP_051618, VSP_051619;
 CC Note=No experimental confirmation available;
 CC TISSUE SPECIFICITY: Blood, predominantly that which is targeted to
 CC the liver.
 CC -1- DOMAIN: Von Willebrand antigen II is required for multimerization
 CC of VWF and for its targeting to storage granules (By similarity).
 CC -1- PTM: All cysteine residues are involved in intrachain or
 CC interchain disulfide bonds (By similarity).
 CC -1- SIMILARITY: Contains 1 CTCK (C-terminal cysteine knot-like) domain.
 CC -1- SIMILARITY: Contains 3 VWF domains.
 CC -1- SIMILARITY: Contains 3 VWF domains.

```

CC -1- SIMILARITY: Contains 4 WFWD domains.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
CC ENBL; AY162409; AAN73055.1; -; mRNA.
CC ENBL; AY208897; AAP41950.1; -; mRNA.
CC ENBL; AF539800; AAN07781.2; -; mRNA.
CC ENBL; AK083237; BAC38822.1; ALT INIT; mRNA.
CC ENBL; AJ238390; CAB86200.1; -; Genomic DNA.
CC ENBL; U27810; AAA82929.1; -; Genomic DNA.
CC PDB; 1U00; X-ray; C-1261-1468.
CC SMR; 1261-1468, 1684-1873.
CC ENsembl; ENSMUSG00000001930; Mus musculus.
CC MGI; MGI:98941; Vwf.
CC GO; GO:0005615; C:extracellular space; TAS.
CC GO; GO:0007596; P:blood coagulation; IMP.
CC InterPro; IPR006207; Cys_knot_C.
CC InterPro; IPR004094; Prot_inh_antistn.
CC InterPro; IPR02919; Prot_inh_CR_TIL.
CC InterPro; IPR008037; Prot_inh_PMP.
CC InterPro; IPR012011; VWF.
CC InterPro; IPR020335; VWF_A.
CC InterPro; IPR001007; VWF_C.
CC InterPro; IPR001846; VWF_D.
CC Pfam; PF02822; Antistatin_1.
CC Pfam; PF05375; Pacifastin_I_1.
CC Pfam; PF01826; TIL; 4.
CC Pfam; PF00092; VWA; 3.
CC Pfam; PF00093; VWC; 3.
CC Pfam; PF00094; VWD; 4.
CC PRINTS; PIRSF002495; VWF; 1.
CC SMART; SM00041; CT; 1.
CC SMART; SM00327; VWA; 3.
CC SMART; SM00214; VWC; 5.
CC SMART; SM00216; VWD; 4.
CC PROSITE; PS01185; CTCK_1; 1.
CC PROSITE; PS01225; CTCK_2; 1.
CC PROSITE; PS0234; VWF; 3.
CC PROSITE; PS01208; VWF; 1; 3.
CC PROSITE; PS01084; VWF; 2; 3.
CC 3D-structure; Alternative splicing; Blood coagulation; Cell adhesion;
KW Extracellular matrix; Glycoprotein; Hemostasis; Repeat; Signal.
FT SIGNAL 1 22 By similarity.
FT CHAIN 23 763 Von Willebrand antigen II.
FT CHAIN 764 2813 Von Willebrand factor.
FT DOMAIN 34 386 WFWD 1.
FT DOMAIN 387 745 WFWD 2.
FT DOMAIN 860 1241 WFWD 3.
FT DOMAIN 1277 1453 WFWD 4.
FT DOMAIN 1498 1665 WFWD 5.
FT DOMAIN 1691 1871 WFWD 6.
FT DOMAIN 1934 2203 WFWD 7.
FT DOMAIN 2255 2328 WFWD 8.
FT DOMAIN 2429 2495 WFWD 9.
FT DOMAIN 2580 2645 WFWD 10.
FT DOMAIN 2724 2812 CTCK.
FT REGION 764 787 Amino-terminal.
FT REGION 788 833 E1.
FT REGION 826 853 E2.
FT MOTIF 2507 2509 Cell attachment site (By similarity).
FT CARBOHYD 99 99 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 156 156 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 666 666 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 857 857 N-linked (GlcNAc...) (By similarity).
FT CARBOHYD 1005 1005 N-linked (GlcNAc...) (Potential).

```

```

FT CARBOHYD 1147 1147 N-linked (GlcNAc...) (By similarity).
FT CARBOHYD 1231 1231 N-linked (GlcNAc...) (By similarity).
FT CARBOHYD 1248 1248 O-linked (GlcNAc...) (By similarity).
FT CARBOHYD 1255 1255 O-linked (GlcNAc...) (By similarity).
FT CARBOHYD 1256 1256 O-linked (GlcNAc...) (By similarity).
FT CARBOHYD 1468 1468 O-linked (GlcNAc...) (By similarity).

Query Match 74.5%; Score 41; DB 1; Length 2813;
Best Local Similarity 77.8%; Pred. No. 1.8e+02; Mismatches 2; Indels 0; Gaps 0;
Matches 7; Conservative 0;

QY 1 WREPSFMA1 9
Db 642 WREPGFCA1 650

RESULT 6
Q9FYA0 ARATH
ID Q9FYA0 ARATH PRELIMINARY; PRT; 354 AA.
AC Q9FYA0;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE Hypothetical protein T19L5_60 (AT5G13100/T19L5_60) (At5g13100).
GN Name=T19L5_60;
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopses.
OX NCBI_TaxID=3702;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Bevan M., Rose M., Hempel S., Entian K.-D., Bancroft I., Mewes H.W.,
RA Rudd S., Lemcke K., Mayer K.F.X.;
RA Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RA EU Arabidopsis sequencing project;
RA Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP NUCLEOTIDE SEQUENCE.
RA Cheuk R., Chen H., Kim C.J., Koesema E., Meyers M.C., Banh J.,
RA Bowser L., Carninci P., Dale J.M., Goldsmith A.D., Hayashizaki Y.,
RA Ishida J., Jiang P.X., Jones T., Kamiya A., Karlin-Neumann G.,
RA Kawai J., Lam B., Lee J.M., Lin J., Liu S.X., Miranda M., Narusaka M.,
RA Nguyen M., Onodera C.S., Palm C.J., Pham P.K., Quach H.L., Sakurai T.,
RA Satou M., Seki M., Southwick A., Tang C.C., Toriumi M., Yamada K.,
RA Yamamura Y., Yu G., Yu S., Shinozaki K., Davis R.W., Theologis A.,
RA Ecker J.R.;
RA Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
RN [4]
RP NUCLEOTIDE SEQUENCE.
RA Cheuk R., Chen H., Kim C.J., Shinn P., Ecker J.R.;
RA "Arabidopsis ORF clones";
RT Submitted (DEC-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL391711; CAC05438.1; -; Genomic DNA.
DR EMBL; AY058168; AAL25582.1; -; mRNA.
DR EMBL; BT020363; AAV85718.1; -; mRNA.
KW Hypothetical protein.
SQ SEQUENCE 354 AA; 39305 MW; 4418A9B4B891138C CRC64;

Query Match 72.7%; Score 40; DB 2; Length 354;
Best Local Similarity 70.0%; Pred. No. 32; Mismatches 1; Indels 0; Gaps 0;
Matches 7; Conservative 2;

QY 1 WREPSFMA1S 10
Db 21 WREPSFMA1S 30

RESULT 7
Q6WH2_9ACTO
ID Q6WH2_9ACTO PRELIMINARY; PRT; 1864 AA.

```

```

AC Q6WHV2;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE EspE.
GN Name=espE;
OS Actinomadura verrucosospora.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptosporangineae; Thermomonosporaceae; Actinomadura.
OX NCBI_TaxID=46165;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=ATCC 39334;
RA MEDLINE=2292109; PubMed=14528002; DOI=10.1073/pnas.2034291100;
RA Liu W., Ahlert J., Gao Q., Wendt-Pienkowski E., Shen B., Thorson J.S.;
RT "Rapid PCR amplification of minimal enediyne polyketide synthase
RL cassettes leads to a predictive familial classification model.";
RL Proc. Natl. Acad. Sci. U.S.A. 100:11959-11963(2003).
DR EMBL; AY267372; AAP22148.1; -; Genomic DNA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0006633; P:fatty acid biosynthesis; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR001227; Ac transferase.
DR InterPro; IPR002198; ADH_short.
DR InterPro; IPR002347; Adh_short_C2.
DR InterPro; IPR000794; Ketoacyl_synth.
DR Pfam; PF00698; Acyl_transf_1; 1.
DR Pfam; PF00106; adh_short; 1.
DR Pfam; PF00109; ketoacyl-synt; 1.
DR Pfam; PF02801; Ketoacyl-synt_C; 1.
DR PRINTS; PR00081; GDRHRD.
SQ SEQUENCE 1864 AA; 193996 MW; 07F72F26381DDC7B CRC64;

Query Match 72.7%; Score 40; DB 2; Length 1864;
Best Local Similarity 75.0%; Pred. No. 1.8e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 WREPSFMA 8
Db 890 WREPEFLA 897

RESULT 8
ID Q9H2F6 PSEAE PRELIMINARY; PRT; 326 AA.
AC Q9H2F6;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein.
GN OrderedLocusNames=PA3052;
OS Pseudomonas aeruginosa.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=287;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=ATCC 15692 / PA01;
RA MEDLINE=20437337; PubMed=10984043; DOI=10.1038/35023079;
RA Stover C.K., Phan X.-Q.T., Erwin A.L., Mizoguchi S.D., Warrenner P.,
RA Hickey M.J., Brinkman P.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
RA Garber R.L., Goltzy L., Tolentino E., Westbrook-Wadman S., Yuan Y.,
RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
RA Smith K.A.H., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
RA Reizer J., Saier M.H. Jr., Hancock R.E.W., Lory S., Olson M.V.;
RT "Complete genome sequence of Pseudomonas aeruginosa PA01, an
RT opportunistic pathogen.";
RL Nature 406:959-964(2000).
DR EMBL; AE004730; AAG06440.1; -; Genomic DNA.
DR PIR; C83263; C83263.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 326 AA; 36362 MW; B63E96D3725A28A6 CRC64;

Query Match 70.9%; Score 39; DB 2; Length 577;
Best Local Similarity 66.7%; Pred. No. 85;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 WREPSFMA 9
Db 265 WSEPTFMSL 273

RESULT 9
ID Q6CED5_YARLI PRELIMINARY; PRT; 577 AA.
AC Q6CED5;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Yarrowia lipolytica chromosome B of strain CLIB99 of Yarrowia
DE lipolytica.
GN OrderedLocusNames=YALI0B16522g;
OS Yarrowia lipolytica (Candida lipolytica).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Dipodascaceae; Yarrowia.
OX NCBI_TaxID=4952;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=CLIB 122 / E 150;
RX PubMed=15229592; DOI=10.1038/nature02579;
RA Dujon B., Sherman D., Fischer G., Durrens P., Casaregola S.,
RA Lafontaine I., de Montigny J., Marck C., Neuvéglise C., Talla E.,
RA Garnay N., Frangeul L., Aigle M., Anthouard V., Babour A., Barbe V.,
RA Barney S., Blanchin S., Beckerich J.-M., Beyne E., Bleykasten C.,
RA Boisrame A., Boyer J., Cattolico L., Confanioli F., de Daruvar A.,
RA Despons L., Fabre E., Fairhead C., Ferry-Dumazet H., Groppi A.,
RA Hantraye F., Hennequin C., Jauniaux N., Joyet P., Kachouri R.,
RA Kerrest A., Koszul R., Lemaire M., Lesur I., Ma L., Muller H.,
RA Nicaud J.-M., Nikolski M., Oztas S., Ozier-Kalogeropoulos O.,
RA Pellenz S., Potier S., Richard G.-F., Straub M.-L., Suleau A.,
RA Swennen D., Tekai F., Wesolowski-Louvel M., Westhof E., Wirth B.,
RA Zeniou-Meyer M., Zivanovic Y., Bolotin-Fukuhara M., Thierry A.,
RA Bouchier C., Caudron B., Scarpelli C., Gaillardin C., Weissenbach J.,
RA Wincker P., Souciet J.-L.;
RT "Genome evolution in yeasts.";
RL Nature 430:35-44(2004).
DR EMBL; CR382128; CAG83230.1; -; Genomic DNA.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0015359; F:amino acid permease activity; IEA.
DR GO; GO:0006865; P:amino acid transport; IEA.
DR InterPro; IPR006810; P:transport; IEA.
DR InterPro; IPR002293; AA/rel_permease.
DR InterPro; IPR004840; AAC_permease.
DR InterPro; IPR004841; Permease region.
DR InterPro; IPR004762; Yeast_AA_perm.
DR Pfam; PF00324; AA_permease; 1.
DR TIGRFAMs; TIGR00913; ZA0310; 1.
DR PROSITE; PS00218; AMINO ACID PERMEASE_1; 1.
KW Complete proteome; Transmembrane; Transport.
SQ SEQUENCE 577 AA; 63096 MW; FF62ADC70DC2B66 CRC64;

Query Match 70.9%; Score 39; DB 2; Length 577;
Best Local Similarity 66.7%; Pred. No. 85;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 WREPSFMA 9
Db 527 WKRPSPFVAL 535

RESULT 10
ID O70535 RAT PRELIMINARY; PRT; 1093 AA.
AC O70535;

```

```

DT 01-AUG-1998 (TrEMBLrel. 07, Created)
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Leukemia inhibitor factor receptor alpha-chain.
GN Name=Liffr;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Wistar-Kimichi; TISSUE=Liver;
RX MEDLINE=98007878; PubMed=9349722; DOI=10.1016/S0167-4781(97)00079-1;
RA Aikawa J., Ikeda-Naiki S., Olgane J., Min K.S., Imamura T., Sasai K.,
RA Shioa K., Ogawa T.;
RT "Molecular cloning of rat leukemia inhibitory factor receptor alpha-
RL Biochim. Biophys. Acta 1353:266-276(1997).
CC -1- DOMAIN: The WSXWS motif appears to be necessary for proper protein
CC folding and thereby efficient intracellular transport and cell-
CC surface receptor binding. (By similarity).
CC -1- DOMAIN: The box 1 motif is required for JAK interaction and/or
CC activation (By similarity).
DR EMBL; D86345; BAA25907.1; -; mRNA.
DR HSSP; P40189; LBOU.
DR Ensembl; ENSRNOG00000001696; Rattus norvegicus.
DR RGD; 621431; Lifr.
DR GO; GO:0016021; C: integral to membrane; IEA.
DR GO; GO:0004872; F: receptor activity; IEA.
DR InterPro; IPR002996; Cytokn_recept_B/G.
DR InterPro; IPR003961; FN III.
DR InterPro; IPR003529; Hemtrecept_1302.
DR Pfam; PF00041; fn3; 4.
DR SMART; SM00060; FN3; 5.
DR PROSITE; PS00853; FN3; 5.
DR PROSITE; PS01353; HEMATOPO_REC_L_F2; UNKNOWN_1.
KW Receptor; Repeat; Transmembrane.
SQ SEQUENCE 1093 AA; 122394 MW; 08D43DEAF8F5E3F6 CRC64;

Query Match 70.9%; Score 39; DB 2; Length 1093;
Best Local Similarity 75.0%; Pred. No. 1.7e+02;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 WREPSPMA 8
Db 7 WRQPSWMA 14

RESULT 11
ID Q8635_HORVU PRELIMINARY; PRT; 48 AA.
AC Q8635;
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein (Fragment).
OS Hordeum vulgare (Barley).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae;
OC Triticeae; Hordeum.
OX NCBI_TaxID=4513;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=albobstrians; TISSUE=White leaf;
RA Hess W.R., Golz R.R., Boerner T.;
RT "Analysis of randomly selected cDNAs reveals the expression of
RT stress- and defence related genes in the barley mutant albobstrians.";
RL Plant Sci. 133:191-201(1998).
DR EMBL; AJ222781; CAA10986.1; -; mRNA.
DR FIR; T05927; T05927.
DR GO; GO:0016020; C: membrane; IEA.
DR InterPro; IPR000537; UbiA_prenyltrans.

```

```

DR Pfam; PF01040; UbiA; 1.
KW Hypothetical protein.
FT NON_TER 1
FT NON_TER 48
SQ SEQUENCE 48 AA; 5351 MW; 1F82FBA71EB35D87 CRC64;

Query Match 69.1%; Score 38; DB 2; Length 48;
Best Local Similarity 60.0%; Pred. No. 9.9;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 WREPSPFMA 10
Db 21 WQSPHPMALA 30

RESULT 12
ID Q5C3K8_SCHJA PRELIMINARY; PRT; 185 AA.
AC Q5C3K8;
DT 10-MAY-2005 (TrEMBLrel. 30, Created)
DT 10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE Hypothetical protein.
OS Schistosoma japonicum (Blood fluke).
OC Eukaryota; Metazoa; Platyhelminthes; Trematoda; Digenea; Strigeidida;
OC Schistosomatoidea; Schistosomatidae; Schistosoma.
OX NCBI_TaxID=6182;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Han Z.;
RL Submitted (MAR-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY809878; AAX25767.1; -; mRNA.
DR InterPro; IPR002563; Flv_red_FWN_bind.
KW Hypothetical protein.
SQ SEQUENCE 185 AA; 19909 MW; A82ABF7CA7CAE103 CRC64;

Query Match 69.1%; Score 38; DB 2; Length 185;
Best Local Similarity 70.0%; Pred. No. 41;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 WREPSPFMA 10
Db 55 WQPPHPMALA 64

RESULT 13
ID Q8YW08_ANASP PRELIMINARY; PRT; 318 AA.
AC Q8YW08;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Alr1808 protein.
GN OrderedLocusNames=alr1808;
OS Anabaena sp. (strain PCC 7120).
OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.
OX NCBI_TaxID=103690;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=21595285; PubMed=11759840;
RA Kaneko T., Nakamura Y., Wolk C.P., Kuritz T., Sasamoto S.,
RA Watanabe A., Iriguchi M., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kohara M., Matsumoto M., Matsuno A., Muraki A.,
RA Nakazaki N., Shimo S., Sugimoto M., Takazawa M., Yamada M.,
RA Yasuda M., Tabata S.;
RT "Complete genomic sequence of the filamentous nitrogen-fixing
RT cyanobacterium Anabaena sp. strain PCC 7120.";
RL DNA Res. 8:205-213(2001).
DR EMBL; BA000019; BAB73507.1; -; Genomic_DNA.
DR FIR; AB2032; AB2032.
DR InterPro; IPR002798; DUF95_TM.
DR Pfam; PF01944; DUF95; 1.
KW Complete proteome.

```

```
SQ SEQUENCE 318 AA; 35735 MW; 517EFCAB7C8EAE6B CRC64;
Query Match 69.1%; Score 38; DB 2; Length 318;
Best Local Similarity 55.6%; Pred. No. 72;
Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Qy 1 WREPSFMA 9
Db 128 WQDPTFMSL 136
|||||:|
|:|:|:|:|

RESULT 14
Q7ZZP1_RANRI PRELIMINARY; PRT; 371 AA.
AC Q7ZZP1;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE GnrH receptor 2
OS Rana ridibunda (laughing frog) (Marsh frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Ranoidea; Ranidae; Rana;
OC Pelophylax.
OX NCBI_TaxID=8406;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Li J.H., Seong J.Y., Kwon H.B.;
RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
CC -|- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
CC -|- SIMILARITY: Belongs to the G-protein coupled receptor 1 family.
DR EMBL: AY260154; AAP15163.1; -; mRNA.
DR GO: GO:0016021; C:integral to membrane; IEA.
DR GO: GO:0004968; F:gonadotropin-releasing hormone receptor act. . .; IEA.
DR GO: GO:0004872; F:receptor activity; IEA.
DR GO: GO:0001584; F:rhodopsin-like receptor activity; IEA.
DR GO: GO:0007186; P:G-protein coupled receptor protein signalin. . .; IEA.
DR GO: GO:0007165; P:signal transduction; IEA.
DR InterPro: IPR000276; GPCR_Rhodopsn.
DR InterPro: IPR001658; GphnRH_receptor.
DR Pfam: PF00001; 7tm 1; 1.
DR PRINTS; PR00529; GNADOTRPHINR.
DR PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR PROSITE; PS0262; G_PROTEIN_RECEP_F1_2; 1.
DR PROSITE; PS00300; SRP54; UNKNOWN 1.
KW G-protein coupled receptor; Receptor; Transducer; Transmembrane.
SQ SEQUENCE 371 AA; 42185 MW; 5CC604568716B267 CRC64;

Query Match 69.1%; Score 38; DB 2; Length 371;
Best Local Similarity 75.0%; Pred. No. 84;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 WREPSFMA 8
Db 29 WREPTFTA 36
|||||:|
|:|:|:|:|

RESULT 15
Q800M0_RANDY PRELIMINARY; PRT; 371 AA.
AC Q800M0;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Gonadotropin-releasing hormone receptor 2.
OS Rana dybowskii (Dybowski's frog) (Korean brown frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Ranoidea; Ranidae; Rana;
OC Rana.
OX NCBI_TaxID=71582;
RN [1]
RP NUCLEOTIDE SEQUENCE.
```

```
RC TISSUE=Brain;
RX MEDLINE=22425989; PubMed=12538604; DOI=10.1210/en.2002-220683;
RA Seong J.Y., Wang L., Oh D.Y., Yun O., Matti K., Li J.H., Soh J.M.,
RA Choi H.S., Kim K., Vaudry H., Kwon H.B.;
RT "Ala/Thr(201) in extracellular loop 2 and Leu/Phe(290) in
RT transmembrane domain 6 of type 1 frog gonadotropin-releasing hormone
RT receptor confer differential ligand sensitivity and signal
RT transduction.";
RL Endocrinology 144:454-466(2003).
CC -|- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
CC -|- SIMILARITY: Belongs to the G-protein coupled receptor 1 family.
DR EMBL: AF236877; AA050196.2; -; mRNA.
DR GO: GO:0016021; C:integral to membrane; IEA.
DR GO: GO:0004968; F:gonadotropin-releasing hormone receptor act. . .; IEA.
DR GO: GO:0004872; F:receptor activity; IEA.
DR GO: GO:0001584; F:rhodopsin-like receptor activity; IEA.
DR GO: GO:0007186; P:G-protein coupled receptor protein signalin. . .; IEA.
DR GO: GO:0007165; P:signal transduction; IEA.
DR InterPro: IPR000276; GPCR_Rhodopsn.
DR InterPro: IPR001658; GphnRH_receptor.
DR Pfam: PF00001; 7tm 1; 1.
DR PRINTS; PR00529; GNADOTRPHINR.
DR PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR PROSITE; PS0262; G_PROTEIN_RECEP_F1_2; 1.
DR PROSITE; PS00300; SRP54; UNKNOWN 1.
KW G-protein coupled receptor; Receptor; Transducer; Transmembrane.
SQ SEQUENCE 371 AA; 42103 MW; 5991880B9BEE2D49 CRC64;
```

```
Query Match 69.1%; Score 38; DB 2; Length 371;
Best Local Similarity 75.0%; Pred. No. 84;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
```

```
Qy 1 WREPSFMA 8
Db 29 WREPTFTA 36
|||||:|
|:|:|:|:|
```

```
Search completed: November 22, 2005, 19:18:54
Job time : 240 secs
```

This Page Blank (uspio)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 22, 2005, 19:08:51 ; Search time 46 Seconds
(without alignments)
17.973 Million cell updates/sec

Title: US-10-733-852-1
Perfect score: 55
Sequence: 1 WREPSFMA1S 10

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/ptodata/1/iaa/5_COMB.pep.*
2: /cgn2_6/ptodata/1/iaa/6_COMB.pep.*
3: /cgn2_6/ptodata/1/iaa/H_COMB.pep.*
4: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep.*
5: /cgn2_6/ptodata/1/iaa/RE_COMB.pep.*
6: /cgn2_6/ptodata/1/iaa/backfilea1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	55	100.0	10	2	US-09-127-134-1
2	55	100.0	10	2	US-09-904-923C-3
3	55	100.0	10	2	US-10-125-332-1
4	55	100.0	16	2	US-08-856-074A-19
5	55	100.0	16	2	US-08-856-074A-22
6	51	92.7	9	1	US-08-470-837-20
7	51	92.7	9	2	US-08-868-452-20
8	51	92.7	9	2	US-08-868-452-51
9	45	81.8	9	2	US-08-868-452-50
10	45	81.8	2813	2	US-08-896-449A-2
11	45	81.8	2813	2	US-09-132-652-2
12	45	81.8	2813	2	US-09-886-900A-2
13	45	81.8	2813	2	US-09-662-478C-2
14	43	78.2	211	2	US-09-489-039A-7326
15	39	70.9	188	2	US-09-252-991A-18496
16	38	69.1	72	2	US-09-328-352-7350
17	38	69.1	124	2	US-09-583-110-4030
18	38	69.1	600	1	US-08-374-155A-10
19	38	69.1	600	1	US-08-785-396-10
20	38	69.1	600	2	US-10-061-269-10
21	38	69.1	629	1	US-08-374-155A-4
22	38	69.1	629	1	US-08-785-396-4
23	38	69.1	629	2	US-10-061-269-4
24	37	-67.3	2050	1	US-08-347-594A-2
25	37	67.3	2813	2	US-09-381-261A-1
26	36	65.5	693	1	US-08-553-279-2
27	35	63.6	304	2	US-09-252-991A-31805

28	35	63.6	311	2	US-09-489-039A-9550	Sequence 9550, Ap
29	35	63.6	431	2	US-09-549-848B-6	Sequence 6, Appli
30	35	63.6	431	2	US-09-688-069-6	Sequence 6, Appli
31	35	63.6	719	1	US-07-943-843-4	Sequence 4, Appli
32	35	63.6	719	1	US-08-347-003-4	Sequence 4, Appli
33	35	63.6	843	2	US-09-252-991A-19886	Sequence 19886, A
34	34	61.8	10	2	US-09-796-149B-1	Sequence 1, Appli
35	34	61.8	502	2	US-09-328-352-6968	Sequence 6968, Ap
36	34	61.8	504	2	US-09-489-039A-8489	Sequence 8489, Ap
37	34	61.8	513	2	US-09-543-681A-8279	Sequence 8279, Ap
38	34	61.8	539	2	US-09-614-912-144	Sequence 144, App
39	33	60.0	108	2	US-09-621-976-7249	Sequence 7249, Ap
40	33	60.0	161	2	US-09-270-767-34626	Sequence 34626, A
41	33	60.0	161	2	US-09-270-767-49843	Sequence 49843, A
42	33	60.0	171	2	US-09-252-991A-21107	Sequence 21107, A
43	33	60.0	220	2	US-09-252-991A-25709	Sequence 25709, A
44	33	60.0	228	2	US-09-603-552-1	Sequence 1, Appli
45	33	60.0	369	2	US-09-489-039A-8107	Sequence 8107, Ap

ALIGNMENTS

RESULT 1
US-09-127-134-1
; Sequence 1, Application US/09127134
; Patent No. 6387663
; GENERAL INFORMATION:
; APPLICANT: Hall, Frederick L
; APPLICANT: Gordon, Erlinda M
; APPLICANT: Starnes, Vaughn A
; APPLICANT: Anderson, W. F
; TITLE OF INVENTION: TARGETING PHARMACEUTICAL AGENTS TO INJURED TISSUES
; FILE REFERENCE: 06666/019001
; CURRENT APPLICATION NUMBER: US/09/127,134
; CURRENT FILING DATE: 1998-07-31
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 1
; TYPE: PRT
; LENGTH: 10
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Artificial
; OTHER INFORMATION: Decapeptide sequence
US-09-127-134-1

Query Match 100.0%; Score 55; DB 2; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.0004;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WREPSFMA1S 10
| | | | | | | | | |
DB 1 WREPSFMA1S 10

RESULT 2
US-09-904-923C-3
; Sequence 3, Application US/09904923C
; Patent No. 6864082
; GENERAL INFORMATION:
; APPLICANT: Anderson, W. French
; APPLICANT: Gordon, Erlinda Maria
; APPLICANT: Hall, Frederick L
; APPLICANT: Starnes, Vaughn A
; TITLE OF INVENTION: Modified viral surface proteins for binding to extracellular matr
; FILE REFERENCE: 271010-463
; CURRENT APPLICATION NUMBER: US/09/904,923C
; CURRENT FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 08/837,223
; PRIOR FILING DATE: 1997-04-10
; NUMBER OF SEQ ID NOS: 8

```
; SOFTWARE: PatentIn version 3.1.1
; SEQ ID NO 3
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: collagen-binding domain of von Willebrand Factor
US-09-904-923C-3

Query Match      100.0%; Score 55; DB 2; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.0004;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 WREPSFMLS 10
Db      1 WREPSFMLS 10

RESULT 3
US-10-125-332-1
; Sequence 1, Application US/10125332
; Patent No. 6955898
; GENERAL INFORMATION:
; APPLICANT: Hall, Frederick L
; APPLICANT: Gordon, Erlinda M
; APPLICANT: Starnes, Vaughn A
; APPLICANT: Anderson, W. F
; TITLE OF INVENTION: TARGETING PHARMACEUTICAL AGENTS TO INJURED TISSUES
; FILE REFERENCE: 06666/019001
; CURRENT APPLICATION NUMBER: US/10/125,332
; CURRENT FILING DATE: 2002-04-17
; PRIOR APPLICATION NUMBER: 09/127,134
; PRIOR FILING DATE: 1998-07-31
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Artificial
; OTHER INFORMATION: Decapeptide sequence
US-10-125-332-1

Query Match      100.0%; Score 55; DB 2; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.0004;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 WREPSFMLS 10
Db      1 WREPSFMLS 10

RESULT 4
US-08-856-074A-19
; Sequence 19, Application US/08856074A
; Patent No. 6004798
; GENERAL INFORMATION:
; APPLICANT: Anderson, W. French
; APPLICANT: Wu, Bonnie W.
; TITLE OF INVENTION: Retroviral Envelopes Having
; TITLE OF INVENTION: Modified Hypervariable polyproline Regions
; NUMBER OF SEQUENCES: 41
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Carella, Byrne, Bain, Gilfillan, Cecchi, Stewart & Olstein
; STREET: 6 Becker Farm Road
; CITY: Roseland
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07068
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch diskette
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: WINDOWS 95
; SOFTWARE: Word 97
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/856,074A
; FILING DATE: 14-May-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Olstein, Elliot M.
; REGISTRATION NUMBER: 24,025
; REFERENCE/DOCKET NUMBER: 271010-378
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 973-994-1700
; TELEFAX: 973-994-1744
; INFORMATION FOR SEQ ID NO: 22:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 16 amino acids

Query Match      100.0%; Score 55; DB 2; Length 16;
Best Local Similarity 100.0%; Pred. No. 0.00068;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 WREPSFMLS 10
Db      4 WREPSFMLS 13

RESULT 5
US-08-856-074A-22
; Sequence 22, Application US/08856074A
; Patent No. 6004798
; GENERAL INFORMATION:
; APPLICANT: Anderson, W. French
; APPLICANT: Wu, Bonnie W.
; TITLE OF INVENTION: Retroviral Envelopes Having
; TITLE OF INVENTION: Modified Hypervariable polyproline Regions
; NUMBER OF SEQUENCES: 41
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Carella, Byrne, Bain, Gilfillan, Cecchi, Stewart & Olstein
; STREET: 6 Becker Farm Road
; CITY: Roseland
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07068
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch diskette
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: WINDOWS 95
; SOFTWARE: Word 97
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/856,074A
; FILING DATE: 14-May-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Olstein, Elliot M.
; REGISTRATION NUMBER: 24,025
; REFERENCE/DOCKET NUMBER: 271010-378
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 973-994-1700
; TELEFAX: 973-994-1744
; INFORMATION FOR SEQ ID NO: 22:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 16 amino acids
```

```
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: polypeptide
US-08-856-074A-22

Query Match 100.0%; Score 55; DB 2; Length 16;
Best Local Similarity 100.0%; Pred. No. 0.00068;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WREPSPFALS 10
Db 4 WREPSPFALS 13

RESULT 6
US-08-470-837-20
; Sequence 20, Application US/08470837
; Patent No. 5808611
; GENERAL INFORMATION:
; APPLICANT: Nimmi, Marcel E.
; APPLICANT: Hall, Frederick L.
; APPLICANT: Tuan, Tai-Lan
; APPLICANT: Wu, Lingtao
; APPLICANT: Cheung, David T.
; TITLE OF INVENTION: Transforming Growth Factor B Fusion
; TITLE OF INVENTION: and
; TITLE OF INVENTION: Their Use in Wound Healing
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Merchant & Gould
; STREET: 11150 Santa Monica Boulevard, Suite 400
; CITY: Los Angeles
; STATE: California
; COUNTRY: USA
; ZIP: 90025-3395
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/470,837
; FILING DATE:
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Sharp, Janice A.
; REGISTRATION NUMBER: 34,051
; REFERENCE/DOCKET NUMBER: 30630-IUS01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 310-445-1140
; TELEFAX: 310-445-9031
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-470-837-20

Query Match 92.7%; Score 51; DB 1; Length 9;
Best Local Similarity 100.0%; Pred. No. 4.6e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WREPSPFAL 9
Db 1 WREPSPFAL 9

RESULT 7
US-08-868-452-20
; Sequence 20, Application US/08868452C
; Patent No. 6352972
; GENERAL INFORMATION:
; APPLICANT: Marcel E. Nimmi
; APPLICANT: Frederick L. Hall
; APPLICANT: Lingtao Wu
; APPLICANT: Bo Han
; APPLICANT: Edwin Shors
; TITLE OF INVENTION: BONE MORPHOGENETIC PROTEINS AND THEIR
; TITLE OF INVENTION: USE IN BONE GROWTH
; FILE REFERENCE: 17972-11
; CURRENT APPLICATION NUMBER: US/08/868,452C
; CURRENT FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 51
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 20
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Human
US-08-868-452-20

Query Match 92.7%; Score 51; DB 2; Length 9;
Best Local Similarity 100.0%; Pred. No. 4.6e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WREPSPFAL 9
Db 1 WREPSPFAL 9

RESULT 8
US-08-868-452-51
; Sequence 51, Application US/08868452C
; Patent No. 6352972
; GENERAL INFORMATION:
; APPLICANT: Marcel E. Nimmi
; APPLICANT: Frederick L. Hall
; APPLICANT: Lingtao Wu
; APPLICANT: Bo Han
; APPLICANT: Edwin Shors
; TITLE OF INVENTION: BONE MORPHOGENETIC PROTEINS AND THEIR
; TITLE OF INVENTION: USE IN BONE GROWTH
; FILE REFERENCE: 17972-11
; CURRENT APPLICATION NUMBER: US/08/868,452C
; CURRENT FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 51
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 51
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Human
US-08-868-452-51

Query Match 92.7%; Score 51; DB 2; Length 9;
Best Local Similarity 100.0%; Pred. No. 4.6e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WREPSPFAL 9
Db 1 WREPSPFAL 9

RESULT 9
US-08-868-452-50
; Sequence 50, Application US/08868452C
; Patent No. 6352972
; GENERAL INFORMATION:
; APPLICANT: Marcel E. Nimmi
; APPLICANT: Frederick L. Hall
; APPLICANT: Lingtao Wu
; APPLICANT: Bo Han
; APPLICANT: Edwin Shors
; TITLE OF INVENTION: BONE MORPHOGENETIC PROTEINS AND THEIR
; TITLE OF INVENTION: USE IN BONE GROWTH
; FILE REFERENCE: 17972-11
```

; CURRENT APPLICATION NUMBER: US/08/868,452C
; CURRENT FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 51
; SOFTWARE: FASTSEQ for Windows Version 3.0
; SEQ ID NO 50
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Human
; US-08-868-452-50

Query Match 81.8%; Score 45; DB 2; Length 9;
Best Local Similarity 88.9%; Pred. No. 4.6e+05;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 WREPSFMA1 9
| | | | | | | | | |
Db 1 WREPSFCA1 9

RESULT 10

US-08-896-449A-2
; Sequence 2, Application US/08896449A
; Patent No. 6040143
; GENERAL INFORMATION:
; APPLICANT: Venta, Patrick J
; APPLICANT: Yuzbasiyan-Gurkan, Vilma
; APPLICANT: Schall, William D
; APPLICANT: Brewer, George J
; TITLE OF INVENTION: DNA ENCODING CANINE VON WILLEBRAND
; TITLE OF INVENTION: FACTOR AND METHODS OF USE
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Harness, Dickey & Pierce, P.L.C.
; STREET: 5445 Corporate Drive
; CITY: Troy
; STATE: Michigan
; COUNTRY: USA
; ZIP: 48098
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION NUMBER: US/08/896,449A
; FILING DATE: 18-JUL-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Smith, DeAnn F.
; REFERENCE/DOCKET NUMBER: 2115-001226
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 248-641-1600
; TELEFAX: 248-641-0270
; TELEX: 287637
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; TYPE: amino acid
; LENGTH: 2813 amino acids
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-896-449A-2

Query Match 81.8%; Score 45; DB 2; Length 2813;
Best Local Similarity 80.0%; Pred. No. 17;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 WREPSFMA1 10
| | | | | | | | | |
Db 642 WREPGFCA1 651

RESULT 11

US-09-132-652-2

; Sequence 2, Application US/09132652
; Patent No. 6074832
; GENERAL INFORMATION:
; APPLICANT: Venta, Patrick J
; APPLICANT: Yuzbasiyan-Gurkan, Vilma
; APPLICANT: Schall, William D
; APPLICANT: Brewer, George J
; APPLICANT: Duffendeck, John
; TITLE OF INVENTION: DNA ENCODING CANINE VON WILLEBRAND FACTOR AND METHODS
; TITLE OF INVENTION: OF USE
; FILE REFERENCE: 21155-001226CPB
; CURRENT APPLICATION NUMBER: US/09/132,652
; CURRENT FILING DATE: 1998-08-11
; EARLIER APPLICATION NUMBER: 08/896,449
; EARLIER FILING DATE: 1997-07-18
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 2813
; TYPE: PRT
; ORGANISM: Canis familiaris
; US-09-132-652-2

Query Match 81.8%; Score 45; DB 2; Length 2813;
Best Local Similarity 80.0%; Pred. No. 17;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 WREPSFMA1 10
| | | | | | | | | |
Db 642 WREPGFCA1 651

RESULT 12

US-09-886-900A-2
; Sequence 2, Application US/09886900A
; Patent No. 6767707
; GENERAL INFORMATION:
; APPLICANT: Venta, Patrick J.
; APPLICANT: Brewer, George J.
; APPLICANT: Vilma, Yuzbasiyan-Gurkan
; APPLICANT: Schall, William D.
; TITLE OF INVENTION: DNA ENCODING CANINE VON WILLEBRAND FACTOR AND METHODS
; TITLE OF INVENTION: OF USE
; FILE REFERENCE: UMW-1226CPCUS
; CURRENT APPLICATION NUMBER: US/09/886,900A
; CURRENT FILING DATE: 2001-06-21
; PRIOR APPLICATION NUMBER: US/09/662,478C
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: 09/132,652
; PRIOR FILING DATE: 1998-08-11
; PRIOR APPLICATION NUMBER: PCT/US99/18153
; PRIOR FILING DATE: 1999-08-10
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 2813
; TYPE: PRT
; ORGANISM: Canis familiaris
; US-09-886-900A-2

Query Match 81.8%; Score 45; DB 2; Length 2813;
Best Local Similarity 80.0%; Pred. No. 17;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 WREPSFMA1 10
| | | | | | | | | |
Db 642 WREPGFCA1 651

RESULT 13

US-09-662-478C-2
; Sequence 2, Application US/09662478C
; Patent No. 6780583

```
; GENERAL INFORMATION:
; APPLICANT: Venta, Patrick J.
; APPLICANT: Brewer, George J.
; APPLICANT: Vilma, Yuzbasiyan-Gurkan
; APPLICANT: Schall, William D.
; TITLE OF INVENTION: DNA ENCODING CANINE VON WILLEBRAND FACTOR AND METHODS
; TITLE OF INVENTION: OF USE
; FILE REFERENCE: UNV-1226CPCUS
; CURRENT APPLICATION NUMBER: US/09/662,478C
; CURRENT FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: 09/132,652
; PRIOR FILING DATE: 1998-08-11
; PRIOR APPLICATION NUMBER: PCT/US99/18153
; PRIOR FILING DATE: 1999-08-10
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 2813
; TYPE: PRT
; ORGANISM: Canis familiaris
; US-09-662-478C-2

Query Match      81.8%; Score 45; DB 2; Length 2813;
Best Local Similarity 80.0%; Pred. No. 17;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy      1 WREPSPFMLS 10
Db      642 WREPFGFALS 651

RESULT 14
US-09-489-039A-7326
; Sequence 7326, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 7326
; LENGTH: 211
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
; US-09-489-039A-7326

Query Match      78.2%; Score 43; DB 2; Length 211;
Best Local Similarity 70.0%; Pred. No. 2.1;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy      1 WREPSPFMLS 10
Db      35 WROPFPMALT 44

RESULT 15
US-09-252-991A-18496
; Sequence 18496, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18

; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 18496
; LENGTH: 188
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (20)
; OTHER INFORMATION: Identity of amino acid at the above locations are unknown.
US-09-252-991A-18496

Query Match      70.9%; Score 39; DB 2; Length 188;
Best Local Similarity 66.7%; Pred. No. 11;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy      1 WREPSPFMLS 9
Db      127 WSEPTFMSL 135

Search completed: November 22, 2005, 19:19:50
Job time : 53 secs
```

Page Blank (uspio)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 22, 2005, 19:11:51 ; Search time 166 Seconds
(without alignments)
25.170 Million cell updates/sec

Title: US-10-733-852-1
Perfect score: 55
Sequence: 1 WREPSFMALS 10

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA_Main:*
1: /cgn2_6/prodata/1/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/prodata/1/pubpaa/US08_PUBCOMB.pep:*
3: /cgn2_6/prodata/1/pubpaa/US09_PUBCOMB.pep:*
4: /cgn2_6/prodata/1/pubpaa/US10A_PUBCOMB.pep:*
5: /cgn2_6/prodata/1/pubpaa/US10B_PUBCOMB.pep:*
6: /cgn2_6/prodata/1/pubpaa/US11_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	55	100.0	10	3	US-09-904-923B-3
2	55	100.0	10	4	US-10-125-332-1
3	55	100.0	10	4	US-10-223-599-7
4	55	100.0	10	4	US-10-404-714-1
5	55	100.0	10	5	US-10-829-926-4
6	55	100.0	10	5	US-10-733-852-1
7	55	100.0	16	4	US-10-223-599-9
8	55	100.0	16	5	US-10-829-926-7
9	49	89.1	15	5	US-10-829-926-1
10	49	89.1	17	5	US-10-733-852-12
11	45	81.8	2813	3	US-09-886-900-2
12	44	80.0	10	4	US-10-374-271-4
13	39	70.9	72	4	US-10-424-599-251936
14	39	70.9	289	4	US-10-424-599-166163
15	38	69.1	43	4	US-10-425-115-198512
16	38	69.1	140	4	US-10-425-115-238424
17	38	69.1	391	4	US-10-369-493-22053
18	38	69.1	468	4	US-10-236-433-14
19	38	69.1	556	4	US-10-437-963-121538
20	38	69.1	600	4	US-10-061-269-10
21	38	69.1	600	4	US-10-463-509-10
22	38	69.1	600	4	US-10-463-569-10
23	38	69.1	629	4	US-10-061-269-4
24	38	69.1	629	4	US-10-463-509-4
25	38	69.1	629	4	US-10-463-569-4
26	37	67.3	336	3	US-09-925-302-459
27	37	67.3	336	3	US-09-925-302-459

28	37	67.3	466	4	US-10-425-114-71188	Sequence 71188, A
29	37	67.3	2050	6	US-11-053-199-1	Sequence 1, Appli
30	37	67.3	2764	5	US-10-741-600-1156	Sequence 1156, Ap
31	37	67.3	2813	3	US-09-381-261A-1	Sequence 1, Appli
32	37	67.3	2813	4	US-10-020-141-4	Sequence 4, Appli
33	37	67.3	2813	4	US-10-021-660-79	Sequence 79, Appli
34	37	67.3	2813	4	US-10-211-462-149	Sequence 149, App
35	37	67.3	2813	5	US-10-741-600-1155	Sequence 1155, Ap
36	37	67.3	2813	6	US-11-043-611-8	Sequence 8, Appli
37	36	65.5	77	4	US-10-425-115-213174	Sequence 213174,
38	36	65.5	87	4	US-10-425-115-338761	Sequence 338761,
39	36	65.5	119	4	US-10-425-115-359923	Sequence 359923,
40	36	65.5	127	4	US-10-425-115-269053	Sequence 269053,
41	36	65.5	151	4	US-10-424-599-248291	Sequence 248291,
42	36	65.5	171	4	US-10-276-774-1486	Sequence 1486, Ap
43	36	65.5	364	5	US-10-450-763-52310	Sequence 52310, A
44	36	65.5	422	4	US-10-437-963-117783	Sequence 117783,
45	36	65.5	475	4	US-10-369-493-11476	Sequence 11476, A

ALIGNMENTS

RESULT 1
US-09-904-923B-3
; Sequence 3, Application US/09904923B
; Publication No. US20030118551A1
; GENERAL INFORMATION:
; APPLICANT: No. US20030118551A1artis AG
; TITLE OF INVENTION: Modified viral surface proteins for binding to extracellular matrix
; FILE REFERENCE: 271010-463
; CURRENT APPLICATION NUMBER: US/09/904,923B
; CURRENT FILING DATE: 2001-07-13
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: Patentn version 3.1
; SEQ ID NO 3
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: collagen-binding domain of von Willebrand Factor
US-09-904-923B-3

Query Match 100.0%; Score 55; DB 3; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.004;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 WREPSFMALS 10
Db 1 WREPSFMALS 10
|||||
|||||

RESULT 2
US-10-125-332-1
; Sequence 1, Application US/10125332
; Publication No. US20020164719A1
; GENERAL INFORMATION:
; APPLICANT: Hall, Frederick L
; APPLICANT: Gordon, Erlinda M
; APPLICANT: Starnes, Vaughn A
; APPLICANT: Anderson, W. F
; TITLE OF INVENTION: TARGETING PHARMACEUTICAL AGENTS TO INJURED TISSUES
; FILE REFERENCE: 06666/019001
; CURRENT APPLICATION NUMBER: US/10/125,332
; CURRENT FILING DATE: 2002-04-17
; PRIOR APPLICATION NUMBER: 09/127,134
; PRIOR FILING DATE: 1998-07-31
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: Patentn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 10
; TYPE: PRT

```
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Artificial
; OTHER INFORMATION: Decapeptide sequence
US-10-125-332-1.

Query Match      100.0%; Score 55; DB 4; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.004;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 WREPSFMLS 10
Db      1 WREPSFMLS 10
|||||

RESULT 3
US-10-223-599-7
; Sequence 7, Application US/10223599
; Publication No. US20030129163A1
; GENERAL INFORMATION:
; APPLICANT: Hall, Frederick L.
; APPLICANT: Gordon, Erlinda Maria
; APPLICANT: Anderson, W. French
; TITLE OF INVENTION: Retroviral Vectors Including Modified Envelope Escort Proteins
; FILE REFERENCE: 4-30666B
; CURRENT APPLICATION NUMBER: US/10/223,599
; CURRENT FILING DATE: 2002-08-19
; PRIOR APPLICATION NUMBER: US 09/698,545
; PRIOR FILING DATE: 2000-10-27
; PRIOR APPLICATION NUMBER: PCT/IB99/00764
; PRIOR FILING DATE: 1999-04-28
; PRIOR APPLICATION NUMBER: US 09/069,398
; PRIOR FILING DATE: 1998-04-29
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 7
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: polypeptide binding region that binds to collagen
US-10-223-599-7

Query Match      100.0%; Score 55; DB 4; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.004;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 WREPSFMLS 10
Db      1 WREPSFMLS 10
|||||

RESULT 4
US-10-404-714-1
; Sequence 1, Application US/10404714
; Publication No. US20030157078A1
; GENERAL INFORMATION:
; APPLICANT: University of Southern California et al.
; TITLE OF INVENTION: IDENTIFICATION OF A PLURIPOTENT
; TITLE OF INVENTION: PRE-MESENCHYMAL, PRE-HEMATOPOIETIC PROGENITOR CELL
; FILE REFERENCE: 06666-041001
; CURRENT APPLICATION NUMBER: US/10/404,714
; CURRENT FILING DATE: 2003-03-31
; PRIOR APPLICATION NUMBER: US/09/619,865
; PRIOR FILING DATE: 2000-07-20
; PRIOR APPLICATION NUMBER: US 60/144,786
; PRIOR FILING DATE: 1999-07-19
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Homosapien
```

```
US-10-404-714-1

Query Match      100.0%; Score 55; DB 4; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.004;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 WREPSFMLS 10
Db      1 WREPSFMLS 10
|||||

RESULT 5
US-10-829-926-4
; Sequence 4, Application US/10829926
; Publication No. US20040253215A1
; GENERAL INFORMATION:
; APPLICANT: Hall, Frederick L.
; APPLICANT: Gordon, Erlinda M.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING
; TITLE OF INVENTION: DISORDERS
; FILE REFERENCE: 14230-004001
; CURRENT APPLICATION NUMBER: US/10/829,926
; CURRENT FILING DATE: 2004-04-21
; PRIOR APPLICATION NUMBER: US 60/464,571
; PRIOR FILING DATE: 2003-04-21
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-829-926-4

Query Match      100.0%; Score 55; DB 5; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.004;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 WREPSFMLS 10
Db      1 WREPSFMLS 10
|||||

RESULT 6
US-10-733-852-1
; Sequence 1, Application US/10733852
; Publication No. US20050037469A1
; GENERAL INFORMATION:
; APPLICANT: Hall, Frederick
; APPLICANT: Nimni, Marcel
; APPLICANT: Beart, Robert W.
; APPLICANT: Gordon, Erlinda M.
; TITLE OF INVENTION: MATRIX-TARGETED FUSION POLYPEPTIDES FOR
; TITLE OF INVENTION: TISSUE REGENERATION AND WOUND HEALING
; FILE REFERENCE: 06666-042001
; CURRENT APPLICATION NUMBER: US/10/733,852
; CURRENT FILING DATE: 2003-12-10
; PRIOR APPLICATION NUMBER: US/09/624,874
; PRIOR FILING DATE: 2001-08-30
; PRIOR APPLICATION NUMBER: 60/145,488
; PRIOR FILING DATE: 1999-07-21
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Bos taurus
US-10-733-852-1

Query Match      100.0%; Score 55; DB 5; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.004;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 WREPSFMLS 10
```



```
Db      1 WREPSFMLS 10
|||||
RESULT 7
US-10-223-599-9
; Sequence 9, Application US/10223599
; Publication No. US20030129163A1
; GENERAL INFORMATION:
; APPLICANT: Hall, Frederick L.
; APPLICANT: Gordon, Erlinda Maria
; APPLICANT: Anderson, W. French
; TITLE OF INVENTION: Retroviral Vectors Including Modified Envelope Escort Proteins
; FILE REFERENCE: 4-30666B
; CURRENT APPLICATION NUMBER: US/10/223,599
; CURRENT FILING DATE: 2002-08-19
; PRIOR APPLICATION NUMBER: US 09/698,545
; PRIOR FILING DATE: 2000-10-27
; PRIOR APPLICATION NUMBER: PCT/IB99/00764
; PRIOR FILING DATE: 1999-04-28
; PRIOR APPLICATION NUMBER: US 09/069,398
; PRIOR FILING DATE: 1998-04-29
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 9
; LENGTH: 16
; TYPE: PRT
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: polypeptide shown in Example 1 including the collagen binding domain
; OTHER INFORMATION: ain and linkers.
US-10-223-599-9

Query Match      100.0%; Score 55; DB 4; Length 16;
Best Local Similarity 100.0%; Pred. No. 0.0062;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 WREPSFMLS 10
Db      4 WREPSFMLS 13
|||||
RESULT 8
US-10-829-926-7
; Sequence 7, Application US/10829926
; Publication No. US20040253215A1
; GENERAL INFORMATION:
; APPLICANT: Hall, Frederick L.
; APPLICANT: Gordon, Erlinda M.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING
; FILE REFERENCE: 14230-004001
; CURRENT APPLICATION NUMBER: US/10/829,926
; CURRENT FILING DATE: 2004-04-21
; PRIOR APPLICATION NUMBER: US 60/464,571
; PRIOR FILING DATE: 2003-04-21
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 16
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetically generated peptide
US-10-829-926-7

Query Match      100.0%; Score 55; DB 5; Length 16;
Best Local Similarity 100.0%; Pred. No. 0.0062;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 WREPSFMLS 10
Db      5 WREPSFMLS 14
|||||
RESULT 9
US-10-829-926-1
; Sequence 1, Application US/10829926
; Publication No. US20040253215A1
; GENERAL INFORMATION:
; APPLICANT: Hall, Frederick L.
; APPLICANT: Gordon, Erlinda M.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING
; FILE REFERENCE: 14230-004001
; CURRENT APPLICATION NUMBER: US/10/829,926
; CURRENT FILING DATE: 2004-04-21
; PRIOR APPLICATION NUMBER: US 60/464,571
; PRIOR FILING DATE: 2003-04-21
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 15
; TYPE: PRT
; ORGANISM: Homo sapiens
; OTHER INFORMATION:
US-10-829-926-1

Query Match      89.1%; Score 49; DB 5; Length 15;
Best Local Similarity 90.0%; Pred. No. 0.072;
Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1 WREPSFMLS 10
Db      4 WREPSFMLS 13
|||||
RESULT 10
US-10-733-852-12
; Sequence 12, Application US/10733852
; Publication No. US20050037469A1
; GENERAL INFORMATION:
; APPLICANT: Hall, Frederick
; APPLICANT: Nimni, Marcel
; APPLICANT: Beart, Robert W.
; APPLICANT: Gordon, Erlinda M.
; TITLE OF INVENTION: MATRIX-TARGETED FUSION POLYPEPTIDES FOR
; FILE REFERENCE: 06666-042001
; CURRENT APPLICATION NUMBER: US/10/733,852
; CURRENT FILING DATE: 2003-12-10
; PRIOR APPLICATION NUMBER: US/09/624,874
; PRIOR FILING DATE: 2001-08-30
; PRIOR APPLICATION NUMBER: 60/145,488
; PRIOR FILING DATE: 1999-07-21
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 17
; TYPE: PRT
; ORGANISM: Bos taurus
; OTHER INFORMATION:
US-10-733-852-12

Query Match      89.1%; Score 49; DB 5; Length 17;
Best Local Similarity 90.0%; Pred. No. 0.081;
Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1 WREPSFMLS 10
Db      4 WREPSFMLS 13
|||||
RESULT 11
US-09-886-900-2
; Sequence 2, Application US/09886900
; Patent No. US20020137051A1
; GENERAL INFORMATION:
```

APPLICANT: Venta, Patrick J
Yuzbasiyan-Gurkan, Vilma
Schall, William D
Brewer, George J
TITLE OF INVENTION: DNA ENCODING CANINE VON WILLEBRAND
FACTOR AND METHODS OF USE
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: Harness, Dickey & Pierce, P.L.C.
STREET: 5445 Corporate Drive
CITY: Troy
STATE: Michigan
COUNTRY: USA
ZIP: 48098
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/886,900
FILING DATE: 21-Jun-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/896,449
FILING DATE: 18-JUL-1997
ATTORNEY/AGENT INFORMATION:
NAME: Smith, DeAnn F.
REFERENCE/DOCKET NUMBER: 2115-001226
TELECOMMUNICATION INFORMATION:
TELEPHONE: 248-641-1600
TELEFAX: 248-641-0270
TELEX: 287637
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 2813 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-886-900-2

Query Match 81.8%; Score 45; DB 3; Length 2813;
Best Local Similarity 80.0%; Pred. No. 54;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 WREPSFMLS 10
DB 642 WREPGFALS 651

RESULT 12
US-10-374-271-4
; Sequence 4, Application US/10374271
; Publication No. US20040167088A1
; GENERAL INFORMATION:
; APPLICANT: WICKHAM, THOMAS J
; APPLICANT: AKIYAMA, MASAKI
; TITLE OF INVENTION: METHOD OF USING ADENOVIRAL VECTORS WITH INCREASED PERSISTENCE IN
; FILE REFERENCE: 220933
; CURRENT APPLICATION NUMBER: US/10/374,271
; CURRENT FILING DATE: 2003-02-25
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Synthetic
US-10-374-271-4

Query Match 80.0%; Score 44; DB 4; Length 10;

Best Local Similarity 80.0%; Pred. No. 0.4;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 WREPSFMLS 10
DB 1 WREPSFMLS 10

RESULT 13
US-10-424-599-251936
; Sequence 251936, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 251936
; LENGTH: 72
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(72)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_69525C.1.pap
US-10-424-599-251936

Query Match 70.9%; Score 39; DB 4; Length 72;
Best Local Similarity 70.0%; Pred. No. 21;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 WREPSFMLS 10
DB 31 WRESTFOALS 40

RESULT 14
US-10-424-599-166163
; Sequence 166163, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 166163
; LENGTH: 289
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_121060C.1.pap
US-10-424-599-166163

Query Match 70.9%; Score 39; DB 4; Length 289;
Best Local Similarity 70.0%; Pred. No. 77;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 WREPSFMLS 10
DB 31 WRESTFOALS 40

Db 245 WRESTFOALS 254

RESULT 15
US-10-425-115-198512
; Sequence 198512, Application US/10425115
; Publication No, US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 198512
; LENGTH: 43
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_112619C.1.pep
US-10-425-115-198512

Query Match 69.1%; Score 38; DB 4; Length 43;
Best Local Similarity 70.0%; Pred. No. 19;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 WREPSFMLS 10
|||
Db 5 WRTPLFMLS 14

Search completed: November 22, 2005, 19:23:20
Job time : 167 secs

Page Blank (usp10)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: \ November 22, 2005, 19:15:02 ; Search time 7 Seconds
(without alignments)
1.614 Million cell updates/sec

Title: US-10-733-852-1

Perfect score: 55

Sequence: 1 WRPSFMA15 10

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 8323 seqs, 1129788 residues

Total number of hits satisfying chosen parameters: 8323

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA New:

- 1: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB_PEP.*
- 2: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB_PEP.*
- 3: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB_PEP.*
- 4: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB_PEP.*
- 5: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB_PEP.*
- 6: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB_PEP.*
- 7: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB_PEP.*
- 8: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB_PEP.*

* Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	55	100.0	10	7	US-11-014-629-3
2	55	100.0	43	1	US-10-816-768-104
3	55	100.0	56	1	US-10-816-768-106
4	55	100.0	64	1	US-10-816-768-108
5	55	100.0	129	1	US-10-816-768-89
6	28	50.9	255	7	US-11-082-389-328
7	28	50.9	475	7	US-11-180-997-4
8	28	50.9	605	1	US-10-689-742-140
9	28	50.9	750	7	US-11-132-285-2
10	28	50.9	1001	7	US-11-132-285-40
11	27.5	50.0	293	1	US-10-131-826A-422
12	27	49.1	247	1	US-10-510-386-242
13	27	49.1	258	7	US-11-074-176-6
14	27	49.1	349	1	US-10-131-826A-424
15	27	49.1	381	1	US-10-689-742-172
16	27	49.1	487	1	US-10-131-826A-528
17	27	49.1	786	1	US-10-467-962B-103
18	27	49.1	941	1	US-10-131-826A-464
19	26.5	48.2	301	7	US-11-074-176-362
20	26.5	48.2	308	7	US-11-074-176-254
21	26	47.3	20	1	US-10-962-145C-16
22	26	47.3	301	1	US-10-131-826A-176
23	26	47.3	334	1	US-10-131-826A-12
24	26	47.3	434	1	US-10-632-150-24
25	26	47.3	514	7	US-11-103-037-3

Sequence 6, Appli
Sequence 60, Appl
Sequence 210, App
Sequence 2, Appli
Sequence 30, Appl
Sequence 31, Appl
Sequence 7, Appli
Sequence 218, App
Sequence 10, Appl
Sequence 216, App
Sequence 35, Appl
Sequence 60, Appl
Sequence 61, Appl
Sequence 8, Appli
Sequence 139, App
Sequence 258, App
Sequence 33, Appl
Sequence 222, App
Sequence 5, Appli
Sequence 1, Appli

26 47.3 515 1 US-10-630-203-6
27 45.5 193 7 US-11-082-389-60
28 45.5 221 1 US-10-510-386-210
29 45.5 240 7 US-11-154-865-2
30 45.5 299 7 US-11-082-389-30
31 45.5 340 7 US-11-109-156-31
32 45.5 400 7 US-11-018-018-7
33 45.5 426 1 US-10-131-826A-218
34 45.5 483 1 US-10-630-203-10
35 45.5 522 7 US-11-022-562-216
36 45.5 622 7 US-11-021-441-35
37 45.5 1857 7 US-11-057-058-60
38 45.5 1857 7 US-11-057-058-61
39 45.5 2504 1 US-10-647-956A-8
40 43.6 19 1 US-10-503-575-139
41 24 43.6 19 1 US-10-503-575-139
42 24 43.6 46 1 US-10-828-033-33
43 24 43.6 51 1 US-10-986-501-222
44 24 43.6 104 7 US-11-065-669-5
45 24 43.6 157 7 US-11-010-954-1

ALIGNMENTS

RESULT 1.
US-11-014-629-3
; Sequence 3, Application US/11014629
; Publication No. US20050244376A1
; GENERAL INFORMATION:
; APPLICANT: Hall, Frederick L.
; APPLICANT: Gordon, Erlinda M.
; TITLE OF INVENTION: TARGETED GENE DELIVERY TO EXTRACELLULAR MATRIX
; FILE REFERENCE: 30863-704.302
; CURRENT APPLICATION NUMBER: US/11/014,629
; CURRENT FILING DATE: 2004-12-15
; PRIOR APPLICATION NUMBER: US 08/837,223
; PRIOR FILING DATE: 1997-04-10
; PRIOR APPLICATION NUMBER: US 09/904,923
; PRIOR FILING DATE: 2001-07-13
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Collagen-binding domain of von Willebrand Factor
US-11-014-629-3

Query Match 100.0%; Score 55; DB 7; Length 10;
Best Local Similarity 100.0%; Pred. No. 5.1e-06;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WRPSFMA15 10
Db 1 WRPSFMA15 10

RESULT 2
US-10-816-768-104
; Sequence 104, Application US/10816768
; Publication No. US20050250936A1
; GENERAL INFORMATION:
; APPLICANT: Oppermann, Hermann
; APPLICANT: Tai, Mei-Sheng
; APPLICANT: McCartney, John
; TITLE OF INVENTION: Modified TGF-beta Superfamily Proteins
; FILE REFERENCE: STK-075
; CURRENT APPLICATION NUMBER: US/10/816,768
; CURRENT FILING DATE: 2004-04-02
; NUMBER OF SEQ ID NOS: 124
; SOFTWARE: PatentIn version 2.0

; SEQ ID NO 104
; LENGTH: 43
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: H2510
US-10-816-768-104

Query Match 100.0%; Score 55; DB 1; Length 43;
Best Local Similarity 100.0%; Pred. No. 2e-05;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 WREPSFMA1S 10
Db 27 WREPSFMA1S 36

RESULT 3

US-10-816-768-106
; Sequence 106, Application US/10816768
; Publication No. US20050250936A1
; GENERAL INFORMATION:
; APPLICANT: Oppermann, Hermann
; APPLICANT: Tai, Mei-Sheng
; APPLICANT: McCartney, John
; TITLE OF INVENTION: Modified TGF-beta Superfamily Proteins
; FILE REFERENCE: STK-075
; CURRENT APPLICATION NUMBER: US/10/816,768
; CURRENT FILING DATE: 2004-04-02
; NUMBER OF SEQ ID NOS: 124
; SOFTWARE: PatentIn version 2.0
; SEQ ID NO 106
; LENGTH: 56
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: H2523
US-10-816-768-106

Query Match 100.0%; Score 55; DB 1; Length 56;
Best Local Similarity 100.0%; Pred. No. 2.6e-05;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 WREPSFMA1S 10
Db 27 WREPSFMA1S 36

RESULT 4

US-10-816-768-108
; Sequence 108, Application US/10816768
; Publication No. US20050250936A1
; GENERAL INFORMATION:
; APPLICANT: Oppermann, Hermann
; APPLICANT: Tai, Mei-Sheng
; APPLICANT: McCartney, John
; TITLE OF INVENTION: Modified TGF-beta Superfamily Proteins
; FILE REFERENCE: STK-075
; CURRENT APPLICATION NUMBER: US/10/816,768
; CURRENT FILING DATE: 2004-04-02
; NUMBER OF SEQ ID NOS: 124
; SOFTWARE: PatentIn version 2.0
; SEQ ID NO 108
; LENGTH: 64
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: H2524
US-10-816-768-108

Query Match 100.0%; Score 55; DB 1; Length 64;
Best Local Similarity 100.0%; Pred. No. 3e-05;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 WREPSFMA1S 10
Db 35 WREPSFMA1S 44

RESULT 5

US-10-816-768-89
; Sequence 89, Application US/10816768
; Publication No. US20050250936A1
; GENERAL INFORMATION:
; APPLICANT: Oppermann, Hermann
; APPLICANT: Tai, Mei-Sheng
; APPLICANT: McCartney, John
; TITLE OF INVENTION: Modified TGF-beta Superfamily Proteins
; FILE REFERENCE: STK-075
; CURRENT APPLICATION NUMBER: US/10/816,768
; CURRENT FILING DATE: 2004-04-02
; NUMBER OF SEQ ID NOS: 124
; SOFTWARE: PatentIn version 2.0
; SEQ ID NO 89
; LENGTH: 129
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: H2487
US-10-816-768-89

Query Match 100.0%; Score 55; DB 1; Length 129;
Best Local Similarity 100.0%; Pred. No. 5.8e-05;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 WREPSFMA1S 10
Db 11 WREPSFMA1S 20

RESULT 6

US-11-082-399-328
; Sequence 328, Application US/11082389
; Publication No. US2005024935A1
; GENERAL INFORMATION:
; APPLICANT: Pompejus, Markus
; APPLICANT: Kroger, Burkhard
; APPLICANT: Schroder, Hartwig
; APPLICANT: Zelder, Oskar
; APPLICANT: Habernauer, Gregor
; TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING PROTEINS INVOLVED IN MEMBRANE SYNTHESIS AND MEMBRANE
; FILE REFERENCE: BGI-131CPCN
; CURRENT APPLICATION NUMBER: US/11/082,389
; CURRENT FILING DATE: 2005-03-16
; PRIOR APPLICATION NUMBER: US 09/603024
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: US 60/141031
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: US 60/143262
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: US 60/151281
; PRIOR FILING DATE: 1999-08-27
; PRIOR APPLICATION NUMBER: DE 19930487.4
; PRIOR FILING DATE: 1999-07-01
; PRIOR APPLICATION NUMBER: DE 19930489.0
; PRIOR FILING DATE: 1999-07-01
; PRIOR APPLICATION NUMBER: DE 19931549.3
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19931550.7
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19932134.5
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19941379.7
; PRIOR FILING DATE: 1999-08-31

; Remaining Prior Application data removed - See File Wrapper or PALM.

; NUMBER OF SEQ ID NOS: 446
; SEQ ID NO 328
; LENGTH: 255
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-11-082-389-328

Query Match 50.9%; Score 28; DB 7; Length 255;
Best Local Similarity 55.6%; Pred. No. 19;
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 WREPSEAL 9
Db 6 WRIPFLMAV 14

RESULT 7
US-11-180-997-4
; Sequence 4, Application US/11180997
; Publication No. US20050250150A1
; GENERAL INFORMATION:
; APPLICANT: Akzo Nobel Patent Department
; TITLE OF INVENTION: Lawsonia Intracellularis Vaccine
; FILE REFERENCE: 2000.605
; CURRENT APPLICATION NUMBER: US/11/180,997
; CURRENT FILING DATE: 2005-07-13
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 4
; LENGTH: 475
; TYPE: PRT
; ORGANISM: Lawsonia intracellularis
US-11-180-997-4

Query Match 50.9%; Score 28; DB 7; Length 475;
Best Local Similarity 57.1%; Pred. No. 34;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 WREPSPM 7
Db 16 WQSSFM 22

RESULT 8
US-10-689-742-140
; Sequence 140, Application US/10689742
; Publication No. US20050250180A1
; GENERAL INFORMATION:
; APPLICANT: Jacobs, Kenneth
; APPLICANT: McCoy, John M
; APPLICANT: Lavallie, Edward R
; APPLICANT: Racie, Lisa A
; APPLICANT: Evans, Cheryl
; APPLICANT: Merberg, David
; APPLICANT: Treacy, Maurice
; APPLICANT: Spaulding, Vikki
; TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES ENCODING THEM
; FILE REFERENCE: 00766.000091.10
; CURRENT APPLICATION NUMBER: US/10/689,742
; CURRENT FILING DATE: 2003-10-22
; PRIOR APPLICATION NUMBER: 09/746,783
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 231
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 140
; LENGTH: 605
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-689-742-140

Query Match 50.9%; Score 28; DB 1; Length 605;
Best Local Similarity 100.0%; Pred. No. 43;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 WREP 4
Db 68 WREP 71

RESULT 9
US-11-132-285-2
; Sequence 2, Application US/11132285
; Publication No. US20050244876A1
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc.
; TITLE OF INVENTION: Human Tumor Necrosis Factor Receptor TR13 and TR14
; FILE REFERENCE: PFS11P1
; CURRENT APPLICATION NUMBER: US/11/132,285
; CURRENT FILING DATE: 2005-05-19
; PRIOR APPLICATION NUMBER: US/10/046,433
; PRIOR FILING DATE: 2002-01-16
; PRIOR APPLICATION NUMBER: 60/261,960
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 09/618,570
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/144,087
; PRIOR FILING DATE: 1999-07-16
; PRIOR APPLICATION NUMBER: 60/149,450
; PRIOR FILING DATE: 1999-07-18
; PRIOR APPLICATION NUMBER: 60/149,712
; PRIOR FILING DATE: 1999-08-20
; PRIOR APPLICATION NUMBER: 60/153,089
; PRIOR FILING DATE: 1999-09-10
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 750
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-132-285-2

Query Match 50.9%; Score 28; DB 7; Length 750;
Best Local Similarity 100.0%; Pred. No. 53;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WREP 4
Db 631 WREP 634

RESULT 10
US-11-132-285-40
; Sequence 40, Application US/11132285
; Publication No. US20050244876A1
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc.
; TITLE OF INVENTION: Human Tumor Necrosis Factor Receptor TR13 and TR14
; FILE REFERENCE: PFS11P1
; CURRENT APPLICATION NUMBER: US/11/132,285
; CURRENT FILING DATE: 2005-05-19
; PRIOR APPLICATION NUMBER: US/10/046,433
; PRIOR FILING DATE: 2002-01-16
; PRIOR APPLICATION NUMBER: 60/261,960
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 09/618,570
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/144,087
; PRIOR FILING DATE: 1999-07-16
; PRIOR APPLICATION NUMBER: 60/149,450
; PRIOR FILING DATE: 1999-07-18
; PRIOR APPLICATION NUMBER: 60/149,712
; PRIOR FILING DATE: 1999-08-20
; PRIOR APPLICATION NUMBER: 60/153,089
; PRIOR FILING DATE: 1999-09-10
; NUMBER OF SEQ ID NOS: 61

```
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 40
; LENGTH: 1001
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-132-285-40

Query Match
Best Local Similarity 50.9%; Score 28; DB 7; Length 1001;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WREP 4
DB 882 WREP 885

RESULT 11
US-10-131-826A-422
; Sequence 422, Application US/10131826A
; Publication No. US20050245730A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerriksen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C128
; CURRENT APPLICATION NUMBER: US/10/131,826A
; CURRENT FILING DATE: 2002-04-24
; PRIOR APPLICATION NUMBER: 60/049911
; PRIOR FILING DATE: 1997-06-18
; PRIOR APPLICATION NUMBER: 60/056974
; PRIOR FILING DATE: 1997-08-26
; PRIOR APPLICATION NUMBER: 60/059113
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059115
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059117
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059122
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059184
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059263
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/059352
; PRIOR FILING DATE: 1997-09-19
; PRIOR APPLICATION NUMBER: 60/059588
; PRIOR FILING DATE: 1997-09-19
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 422
; LENGTH: 293
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-131-826A-422

Query Match
Best Local Similarity 50.0%; Score 27.5; DB 1; Length 293;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WREP 4
DB 882 WREP 885

RESULT 12
US-10-510-386-242
; Sequence 242, Application US/10510386
; Publication No. US2005024922A1
; GENERAL INFORMATION:
; APPLICANT: Andersen, Jens Tonne
; APPLICANT: Clausen, Ib Groth
; APPLICANT: Jorgensen, Steen Troels
; APPLICANT: Olsen, Peter Bjarke
; APPLICANT: Rasmussen, Michael Dolberg
; TITLE OF INVENTION: Improved Bacillus Host Cell
; FILE REFERENCE: 10294-204-US
; CURRENT APPLICATION NUMBER: US/10/510,386
; CURRENT FILING DATE: 2004-10-04
; NUMBER OF SEQ ID NOS: 248
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 242
; LENGTH: 247
; TYPE: PRT
; ORGANISM: Bacillus licheniformis
US-10-510-386-242

Query Match
Best Local Similarity 49.1%; Score 27; DB 1; Length 247;
Matches 4; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 3 EPSFMALS 10
DB 195 EPNFVSIS 202

RESULT 13
US-11-074-176-6
; Sequence 6, Application US/11074176
; Publication No. US20050250135A1
; GENERAL INFORMATION:
; APPLICANT: Klaenhammer, Todd R.
; APPLICANT: Russell, William M.
; APPLICANT: Altermann, Eric
; APPLICANT: McAuliffe, Olivia
; APPLICANT: Peril, Andrea Azcarate
; TITLE OF INVENTION: Nucleic Acid Sequences Encoding
; FILE REFERENCE: 5051-694
; CURRENT APPLICATION NUMBER: US/11/074,176
; CURRENT FILING DATE: 2005-03-07
; PRIOR APPLICATION NUMBER: 60/551,161
; PRIOR FILING DATE: 2004-03-08
; NUMBER OF SEQ ID NOS: 381
; SOFTWARE: PastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 298
; TYPE: PRT
; ORGANISM: Lactobacillus acidophilus
US-11-074-176-6

Query Match
Best Local Similarity 49.1%; Score 27; DB 7; Length 298;
Matches 4; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 WREP 4
DB 58 WPNPAINMS 67

RESULT 14
```


US-10-131-826A-424
; Sequence 424, Application US/10131826A
; Publication No. US20050245730A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C128
; CURRENT APPLICATION NUMBER: US/10/131,826A
; CURRENT FILING DATE: 2002-04-24
; PRIOR APPLICATION NUMBER: 60/049911
; PRIOR FILING DATE: 1997-06-18
; PRIOR APPLICATION NUMBER: 60/056974
; PRIOR FILING DATE: 1997-08-26
; PRIOR APPLICATION NUMBER: 60/059113
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059115
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059117
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059122
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059184
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059263
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/059352
; PRIOR FILING DATE: 1997-09-19
; PRIOR APPLICATION NUMBER: 60/059588
; PRIOR FILING DATE: 1997-09-19
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 424
; LENGTH: 349
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-131-826A-424

Query Match 49.1%; Score 27; DB 1; Length 349;
Best Local Similarity 30.0%; Pred. No. 40;
Matches 3; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 WREPSFMLS 10
| : | : | : |
Db 149 WKSPASLTMS 158

RESULT 15
US-10-689-742-172
; Sequence 172, Application US/10689742
; Publication No. US20050250180A1
; GENERAL INFORMATION:
; APPLICANT: Jacobs, Kenneth
; APPLICANT: McCoy, John M
; APPLICANT: Lavallie, Edward R
; APPLICANT: Racie, Lisa A
; APPLICANT: Evans, Cheryl

; APPLICANT: Merberg, David
; APPLICANT: Treacy, Maurice
; APPLICANT: Spaulding, Vikki
; TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES ENCODING THEM
; FILE REFERENCE: 00766.000091.10
; CURRENT APPLICATION NUMBER: US/10/689,742
; CURRENT FILING DATE: 2003-10-22
; PRIOR APPLICATION NUMBER: 09/746,783
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 231
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 172
; LENGTH: 381
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-689-742-172

Query Match 49.1%; Score 27; DB 1; Length 381;
Best Local Similarity 57.1%; Pred. No. 43;
Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 WREPSFM 7
| : | : | : |
Db 346 WPEPSLL 352

Search completed: November 22, 2005, 19:23:33
Job time : 8 secs

This Page Blank (uspio)